

# Nonequilibrium fluctuation response relation in a time scale separated system

Shou-Wen Wang,<sup>1,2</sup> Kyogo Kawaguchi,<sup>3</sup> Shin-ichi Sasa,<sup>4</sup> and Lei-Han Tang<sup>1,5</sup>

<sup>1</sup>*Beijing Computational Science Research Center, Beijing, 100094, China*

<sup>2</sup>*Department of Engineering Physics, Tsinghua University, Beijing, 100086, China*

<sup>3</sup>*Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA*

<sup>4</sup>*Department of Physics, Kyoto University, Kyoto 606-8502, Japan*

<sup>5</sup>*Department of Physics and Institute of Computational and Theoretical Studies,  
Hong Kong Baptist University, Hong Kong, China*

(Dated: October 4, 2016)

We present a theoretical framework to analyze the violation of *fluctuation-response relation* (FRR) for any observable from a finite Markov system with two well-separated time scales. We find that, generally for both slow and fast observables, a broad plateau exists in the intermediate frequency region, which contributes to a finite hidden entropy production. Assuming that non-equilibrium behavior arises only from coupling of slow and fast processes, we find that, at large observation time scale, the effective temperature for a slow observable deviates only slightly from the bath temperature, accompanied by an emerging well-defined effective potential landscape, while the deviation is significant for a fast observable. Our study also identifies a wider range of applicability of the Harada-Sasa equality in Markov jumping systems.

## I. INTRODUCTION

In the past three decades, much efforts have been devoted to understanding non-equilibrium thermodynamics, focusing especially on mesoscopic scale non-equilibrium phenomena, where fluctuations are strong [1, 2]. These theoretical studies have led to interesting applications in biological systems, including kinetical proof-reading [3, 4], environment sensing and adaptation [5–7], and oscillation maintenance within cells [8], efficiency of nanosized rotary motors [9, 10]. Due to the intrinsic complexity of these molecular devices and machines, coarse-graining is inevitable to simplify the description.

Recently, there have been much interest to consider the effect of coarse-graining on entropy production of the system [11–19], which is a key thermodynamic quantity that measures how far the system is away from equilibrium. For a system with two distinct time scales, coarse-graining is achieved by adiabatically eliminating the fast variables to obtain a simpler description in terms of the slow variables. Hondou *et al.* first realized that, for a Brownian particle moving in a spatially modulated temperature field, over-damped approximation gives the wrong entropy production rate compared with the under-damped Langevin description, which implies that Carnot efficiency cannot be achieved in such a Brownian heat engine [20, 21]. Esposito made a systematic study on Markov systems with two time scales and found that the total entropy production can be decomposed into that at the coarse-grained level, that only due to the fast dynamics, and that due to the coupling between slow and fast dynamics. Surprisingly, the coupling term is non-negative and thus cannot be ignored in general [22]. The missing contribution at the coarse-grained level is referred to as “hidden entropy production” by Kawaguchi *et al.*, who further showed that it satisfies fluctuation theorem [23]. In our very recent paper, we showed that hidden entropy production reveals itself as a characteristic

plateau in the violation spectrum of *fluctuation-response relation* (FRR) of a slow observable [24]. Our discovery suggests a way to reveal the hidden fast dissipative processes by just studying the trajectory of a slow variable with sufficient temporal resolution.

FRR is a fundamental property of an equilibrium system [25], and its violation can be exploited to characterize non-equilibrium system. This has been applied to study active hair bundles [26], active cytoskeletal networks [27], and molecular motors [28]. Furthermore, one may introduce effective temperature as the ratio between correlation and response, and use its deviation from bath temperature to quantify deviation of the system from equilibrium. Although initially proposed by Cugliandolo *et al.* to characterize glassy systems [29], it has been used recently for small molecules driven out of equilibrium [30]. A more foundational connection between FRR violation and dissipation was pointed out by Harada and Sasa a decade ago in the context of general Langevin systems [31, 32]. Referred to as the Harada-Sasa equality, this relation has been confirmed experimentally in a driven colloid system, and has also been used to study the energetics of F1 ATPase [33], a rotary motor with much higher complexity [28]. Although there is no general connection between FRR violation and dissipation in discrete Markov systems, Lippiello *et al.* generalized the Harada-Sasa equality to *diffusive* Markov jumping systems where the entropy production in the medium for each jump is relatively small [34]. However, the requirement for this generalization is still not very clear.

In this paper, we systematically discuss the FRR violation of a Markov system with two distinct time scales and a finite state space. Its FRR violation spectrum for both a slow and a fast observable is derived, and the connection to hidden entropy production and effective temperature is also discussed. The paper is organized as follows. Section II gives a brief introduction to Harada-Sasa equality. Its generalization to Markov jumping systems

is discussed in Section III. In Section IV, we derive the analytical forms of correlation and response spectrum for Markov systems. Section V introduces our finite Markov model with two time scales, followed by a perturbative analysis of this system, and its FRR violation spectrum for fast and slow observables, respectively. In Section VI, the connection to entropy production partition and to effective temperature are discussed. Section VII illustrates our main idea through an example of sensory adaptation in *E.coli*. We conclude in Section VIII.

## II. THE HARADA-SASA EQUALITY

The FRR violation of a specific degree of freedom can be related to the dissipation rate of the same degree of freedom, as exemplified by the Harada-Sasa equality in the context of Langevin systems. Consider the following  $N_0$ -component Langevin equation

$$\gamma_j \dot{x}_j = F_j(\vec{x}(t), t) + \xi_j(t) + h_j(t), \quad (1)$$

where  $\gamma_j$  is the friction coefficient for the variable  $x_j$ ,  $F_j$  is a driving force that depends on the system configuration  $\vec{x} = (x_1, x_2, \dots)$  and external driving,  $\xi(t)$  the zero-mean white Gaussian noise with variance  $2\gamma_j T$ . The Boltzmann constant  $k_B$  is set to be 1 throughout this article. We assume that the external driving is such that the system reaches a NESS in the time scale much larger than the characteristic operation time.  $h_j(t)$  is a perturbative force that is applied only when we want to measure the linear response function of the system, defined to be  $R_{\dot{x}_j}(t - \tau) \equiv \delta \langle \dot{x}_j(t) \rangle / \delta h_j(\tau)$ . Then, the average heat dissipation rate through the frictional motion of  $x_j$  is given by  $J_j \equiv \langle [\gamma_j \dot{x}_j(t) - \xi_j(t)] \circ \dot{x}_j(t) \rangle_{ss}$ , where  $\circ$  denotes the Stratonovich integral [35] and  $\langle \cdot \rangle_{ss}$  denotes the steady state ensemble. In this general Langevin setting, the Harada-Sasa equality states that [31, 32]

$$J_j = \gamma_j \left\{ \langle \dot{x}_j \rangle_{ss}^2 + \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} [\tilde{C}_{\dot{x}_j}(\omega) - 2T \tilde{R}'_{\dot{x}_j}(\omega)] \right\}. \quad (2)$$

Here, the prime denotes the real part,  $\tilde{R}_{\dot{x}_j}(\omega)$  is the Fourier transform of the response function  $R_{\dot{x}_j}(t - \tau)$ , and  $\tilde{C}_{\dot{x}_j}(\omega)$  is the Fourier transform of the correlation function  $C_{\dot{x}_j}(t - \tau) \equiv \langle [\dot{x}_j(t) - \langle \dot{x}_j \rangle_{ss}] [\dot{x}_j(\tau) - \langle \dot{x}_j \rangle_{ss}] \rangle_{ss}$ . The Fourier transform for a general function  $g(t)$  is defined to be  $\tilde{g}(\omega) \equiv \int_{-\infty}^{\infty} g(t) \exp(i\omega t) dt$ , with  $i$  being the imaginary unit. In the special case of equilibrium,  $\tilde{C}_{\dot{x}_j}(\omega) = 2T \tilde{R}'_{\dot{x}_j}(\omega)$  due to FRR, and the mean drift  $\langle \dot{x}_j \rangle_{ss}$  and heat dissipation rate  $J_j$  also vanishes. In the steady state, the total entropy production rate  $\sigma$  of the system is related to the heat dissipation rate through

$$\sigma = \frac{1}{T} \sum_j J_j. \quad (3)$$

This is the basis for estimating entropy production rate through analyzing the FRR violation for each channel  $x_j$ .

## III. GENERALIZING HARADA-SASA EQUALITY

So far, the condition under which the generalized Harada-Sasa equality holds in Markov jumping systems is still not clear. Here, we present a systematic derivation of the generalized Harada-Sasa equality Eq. (8) and Eq. (18), and identify their range of applicability.

### A. General equality concerning FRR violation

Consider a general Markov process. The transition from state  $n$  to state  $m$  happens with rate  $w_n^m$ . We assume that if  $w_n^m > 0$ , then the reverse rate  $w_m^n > 0$ . The self-transition is prohibited, i.e.,  $w_n^n = 0$ . The probability  $P_n(t)$  at state  $n$  and time  $t$  evolves according to the following master equation

$$\frac{d}{dt} P_n(t) = \sum_m M_{nm} P_m(t), \quad (4)$$

where  $M$  is assumed to be an irreducible transition rate matrix determined by  $M_{nm} = w_n^m - \delta_{nm} \sum_k w_n^k$ . We consider that an external perturbation  $h$  modifies the transition rate in the following way

$$\tilde{w}_m^n = w_m^n \exp \left[ h \frac{\mathcal{Q}_n - \mathcal{Q}_m}{2T} \right], \quad (5)$$

which is a generalization of the way Langevin system is perturbed [36, 37]. Here,  $\mathcal{Q}_m$  is a conjugate variable to perturbation  $h$ .

The linear response of an arbitrary observable  $A$  for this Markov system is defined to be  $R_A(t - \tau) \equiv \delta \langle A(t) \rangle / \delta h(\tau)$ . In the last decade, much efforts have been devoted to study the relation between linear response and fluctuation in non-equilibrium steady state. By using path-integral approach, Baiesi *et al.* have derived the following relation [38]:

$$R_A(t - \tau) = -\frac{\beta}{2} \langle \bar{v}(\tau) A(t) \rangle_{ss} + \frac{\beta}{2} \langle \dot{Q}(\tau) A(t) \rangle_{ss}. \quad (6)$$

Here, along a stochastic trajectory  $n_t$ ,  $Q(t) \equiv \mathcal{Q}_{n_t}$  and  $\dot{Q}(t)$  is the corresponding instantaneous change rate of  $Q(t)$ . We define

$$\bar{v}_n \equiv \sum_{n'} \omega_n^{n'} (\mathcal{Q}_{n'} - \mathcal{Q}_n), \quad (7)$$

which measures the average change rate of  $Q(t)$  conditioned at the initial state  $n$ . Then,  $\bar{v}(\tau) \equiv \bar{v}_{n_\tau}$ . For equilibrium systems,  $\langle \dot{Q}(\tau) A(t) \rangle_{eq} = -\langle \bar{v}(\tau) A(t) \rangle_{eq}$  when  $t > \tau$  and  $\langle \dot{Q}(\tau) A(t) \rangle_{eq} = \langle \bar{v}(\tau) A(t) \rangle_{eq}$  when  $t < \tau$ . These relations reduce Eq. (6) to FRR in equilibrium.

Now, we focus on a specific application of Eq. (6) by choosing the observable  $A(t)$  the same as  $\dot{Q}(t)$  and setting

$t = \tau$  to consider the immediate response  $R_{\dot{Q}}(0)$ . After a little rearrangement, we derive

$$\langle \bar{v}(t) \dot{Q}(t) \rangle_{ss} = \langle \dot{Q} \rangle_{ss}^2 + \int_{-\infty}^{\infty} [\tilde{C}_{\dot{Q}}(\omega) - 2k_B T \tilde{R}'_{\dot{Q}}(\omega)] \frac{d\omega}{2\pi}, \quad (8)$$

where the auto-correlation function  $C_{\dot{Q}}(t-\tau) \equiv \langle (\dot{Q}(t) - \langle \dot{Q} \rangle_{ss})(\dot{Q}(\tau) - \langle \dot{Q} \rangle_{ss}) \rangle_{ss}$ . Assuming that the system jumps from state  $n_{\tau_j^-}$  to  $n_{\tau_j^+}$  at the transition time  $\tau_j$ ,  $\dot{Q}(t) = \sum_j \delta(t - \tau_j) [\mathcal{Q}_{n_{\tau_j^+}} - \mathcal{Q}_{n_{\tau_j^-}}]$ , which takes non-zero value only at the transition time  $\tau_j$ . However, the observable  $\bar{v}(t) \equiv \bar{v}_{n_t}$  is not-well defined at the time of transition. This makes the evaluation of the correlation  $\langle \bar{v}(t) \dot{Q}(t) \rangle$  nontrivial. Here, we define

$$\bar{v}(t) \equiv \frac{1}{2} [\bar{v}_{n_{t+}} + \bar{v}_{n_{t-}}], \quad (9)$$

which takes the medium value at the transition. To evaluate  $\langle \bar{v}(t) \dot{Q}(t) \rangle$ , we only need to consider the transition events. For an ensemble of transition from state  $n$  to  $m$ ,  $\bar{v}(t)$  gives  $[\bar{v}_n + \bar{v}_m]/2$ , while  $\dot{Q}(t)$  gives  $P_n w_n^m (\mathcal{Q}_n - \mathcal{Q}_m)$ , with  $P_n w_n^m$  the average rate for such transition to occur. Then, summing over all possible transitions, we derive

$$\langle \bar{v}(t) \dot{Q}(t) \rangle_{ss} = \frac{1}{4} \sum_{n,m} (\bar{v}_n + \bar{v}_m) [P_m \omega_m^n - P_n \omega_n^m] (\mathcal{Q}_n - \mathcal{Q}_m), \quad (10)$$

where we have symmetrized the result which gives rise to an additional factor 1/2. Here,  $\langle \bar{v}(t) \dot{Q}(t) \rangle_{ss}$  is proportional to the net flux  $P_m \omega_m^n - P_n \omega_n^m$ , which vanishes at equilibrium, in agreement with our expectation. We claim original contribution for the derivation of Eq. (8) and Eq. (10). Note that while the rhs of Eq. (8) is very similar to that of the Harada-Sasa equality, the lhs of Eq. (8) cannot in general be interpreted as heat dissipation for the degree of freedom  $Q(t)$ .

### B. Connection FRR violation to heat dissipation rate

Here, we justify our medium value interpretation concerning Eq. (9) by referring to Langevin cases. Consider the observable  $\dot{x}_j(t)$  for the Langevin equation (1). Here, for simplicity we assume that the force does not have explicit time-dependence, and the NESS is achieved by breaking detailed balance in other ways. The average change rate for  $x_j$  at the position  $\vec{x}$  is given by  $\bar{v}_{\vec{x}} = F_j(\vec{x})/\gamma_j$ . For a transition from  $\vec{x}$  to  $\vec{x}' = \vec{x} + 2\delta\vec{x}$ , our medium value interpretation implies that

$$\bar{v} = \frac{F_j(\vec{x}) + F_j(\vec{x}')}{2\gamma_j} = \frac{1}{\gamma_j} F_j \left( \frac{\vec{x} + \vec{x}'}{2} \right) + O(\delta\vec{x}^2). \quad (11)$$

The second line is obtained by Talyor expansion. For a trajectory with update time  $\delta t$ ,  $\delta\vec{x}^2 \sim \delta t$  and  $\dot{x}_j \sim$

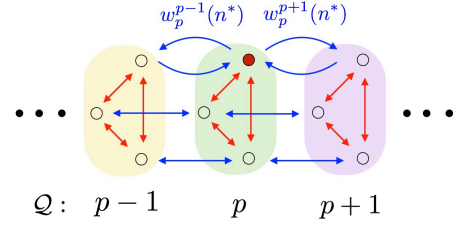


FIG. 1. Illustration of a multi-dimensional hopping process. The black dots represents remaining blocks.

$\delta t^{-1/2}$ . If we use the temporal average to approximate the ensemble average, we have

$$\langle \bar{v}(t) \dot{x}_j(t) \rangle_{ss} = \frac{1}{\gamma_j} \langle F_j(\vec{x}(t)) \circ \dot{x}_j(t) \rangle_{ss} + O(\sqrt{\delta t}). \quad (12)$$

Therefore, our medium value interpretation recovers the Stratonovich interpretation in the limit of  $\delta t \rightarrow 0$ , and Eq. (8) is reduced to the Harada-Sasa equality Eq. (2).

This result can be easily generalized to a multi-dimensional hopping process. The network should be easy to be decomposed into different directions. Let us take FIG. 1 as an example, where the transition could be decomposed into the inter-block direction, indicated in blue arrows, and intra-block direction, labeled in red. The three-state block represents a more complicated sub-network that can also be decomposed into different directions.

Here, we consider first how to capture the dissipation induced by the blue transitions. The key is to choose a proper observable  $\dot{Q}(t)$  which could just count the blue transitions with a proper weight  $\bar{v}(t)$  that relates to the dissipation during the corresponding transition. To do so, we set the conjugate variable  $Q$  to be uniform within each block and make it change value by 1 when jumping to a neighboring block. Let  $n^*$  represents a label for states inside the block. For the state  $(p, n^*)$ , which is labeled in red in FIG. 1, the average change rate for  $Q$  becomes

$$\bar{v}_{p,n^*} = w_p^{p+1}(n^*) - w_p^{p-1}(n^*) \quad (13)$$

according to Eq. (7), which gives the inherent property of this state. Similar to the above discussion of Langevin processes,  $\bar{v}_{p,n^*}$  can be related to some kind of force, or dissipation per jump. We denote

$$\Delta S(n^*, p) \equiv \ln w_p^{p+1}(n^*) / w_{p+1}^p(n^*) \quad (14)$$

as the entropy produced in the medium during the transition from state  $(n^*, p)$  to  $(n^*, p+1)$ . Then  $T\Delta S(n^*, p)$  gives the heat dissipation for this jump. Now, if we assume that for all the blue transitions in FIG. 1 the transition rate satisfies the form

$$w_p^{p+1}(n^*) = \frac{1}{\tau_Q} \exp \left( (1 - \theta) \Delta S(n^*, p) \right) \quad (15a)$$

$$w_{p+1}^p(n^*) = \frac{1}{\tau_Q} \exp \left( -\theta \Delta S(n^*, p) \right), \quad (15b)$$

where  $\tau_Q$  is a constant number and  $\theta$  is a load sharing factor. Then,

$$\bar{\nu}_{p,n^*} = \frac{1}{\tau_Q} (\Delta S(p) + \theta[\Delta S(p-1) - \Delta S(p)] + O(\Delta S^2)), \quad (16)$$

where we have suppressed for  $\Delta S$  the dependence on state  $(n^*, p)$  for simplicity. Assuming that  $|\Delta S| \ll 1$  and that  $\Delta S$  varies slowly along  $p$ , i.e.,  $[\Delta S(p-1) - \Delta S(p)] \ll 1$ , we now successfully connect the average change rate of  $\mathcal{Q}$  with the dissipation per jump:

$$\bar{\nu}_{p,n^*} \approx \frac{1}{\tau_Q} \Delta S(p, n^*). \quad (17)$$

Combined with Eq. (10) and Eq. (14), the integral of the FRR violation for this observable is given by

$$\langle \bar{\nu}(t) \dot{Q}(t) \rangle_{ss} \approx \frac{1}{\tau_Q} \sum_{n^*, p} (P_p^{ss} w_p^{p+1} - P_{p+1}^{ss} w_{p+1}^p) \ln \frac{w_p^{p+1}}{w_{p+1}^p}. \quad (18)$$

Therefore, combined with Eq. (8), this equation implies that the FRR violation of the observable  $\dot{Q}$  captures the dissipation rate due to the inter-block transitions, as indicated by the blue arrows. This equation along with Eq. (8) constitute our generalized Harada-Sasa equality. Here,  $T\tau_Q$  is the corresponding friction coefficient  $\gamma_Q$ .

Now, we make some comments related to Eq. (18).

(a) Key assumptions include that all the inter-block transitions share the same timescale  $\tau_Q$  specified by Eq. (15), that the dissipation per jump is relatively small compared with the thermal energy  $T$ , and that the dissipation changes slowly for neighboring jumps between blocks. However, the load sharing factor  $\theta$  is not required to be 1/2, which was assumed previously by Lippiello *et al.* [34].

(b) The observable  $Q$  is chosen to be a linear function along the block hopping direction such that  $\dot{Q}(t)$  would count the transitions. Even if the actual observable in the experiment is not such a linear form given in FIG. 1, we can map the observed trajectory  $\dot{Q}_{old}$  to a new one  $\dot{Q}_{new}$  associated with a properly designed observable at the stage of data analysis.

(c) To access the dissipation rates due to the inter-block transitions, we have made no assumptions about the transitions inside the blocks. However, in order to access the total dissipation rates, we need to devise other observables to count the transitions inside the blocks and these transitions should satisfy similar constraints.

(d) In certain cases, although  $\Delta S$  is not always small, the probability flux becomes dominant only around the transitions where  $\Delta S$  is small. Therefore, Eq. (18) may also be valid, as illustrated later by our example in FIG. 8.

## IV. CORRELATION AND RESPONSE IN A GENERAL MARKOV SYSTEM

### A. Setup

Here, assuming general Markov processes, we derive the velocity correlation spectrum Eq. (21) and response spectrum Eq. (21) for a general observable, and its FRR violation spectrum Eq. (28). Our strategy is to project these spectrum in the eigenspace of the evolution operator.

Consider a general Markov process the similar as introduced above, except that it has only finite states, say  $N$  states. The  $j$ -th left and right eigenmodes, denoted as  $x_j(n)$  and  $y_j(n)$  respectively, satisfy the eigenvalue equation

$$\sum_m M_{nm} x_j(m) = -\lambda_j x_j(n) \quad (19a)$$

$$\sum_m y_j(m) M_{mn} = -\lambda_j y_j(n), \quad (19b)$$

where the minus sign is introduced to have a positive “eigenvalue”  $\lambda_j$  [39]. These eigenvalues are arranged in the ascending order by their real part, i.e.,  $\text{Re}(\lambda_1) \leq \text{Re}(\lambda_2) \leq \dots$ . This system will reach a unique stationary state associated with  $\lambda_1 = 0$ , where  $y_1 = 1$  and  $x_1(m) = P_m^{ss}$  due to probability conservation. With the proper normalization  $\sum_m P_m^{ss} = 1$ , the eigenmodes satisfy the orthogonal relations  $\sum_m x_j(m) y_{j'}(m) = \delta_{jj'}$  and completeness relations  $\sum_j x_j(n) y_j(m) = \delta_{nm}$ .

### B. Correlation spectrum

Consider first the auto correlation function  $C_Q(t - \tau)$  for the observable  $Q(t) \equiv \mathcal{Q}_{n_t}$ , which can be reformulated in the following form

$$C_Q(t - \tau) = \sum_{n, n'} \mathcal{Q}_n \mathcal{Q}_{n'} P(t - \tau; n, n') P_{n'}^{ss} - \langle Q \rangle_{ss}^2,$$

where  $P(t; n, n')$  is the probability that a system starting in state  $n'$  at time  $t = \tau$  would reach state  $n$  at time  $t$ . An expansion in the eigen space gives  $P(t; n, n') = \sum_j y_j(n') e^{-\lambda_j |t - \tau|} x_j(n)$ , which satisfies the master equation Eq. (4) and the initial condition  $P(0; n, n') = \delta_{nn'}$ . Introducing the weighted average of  $\mathcal{Q}$  in the  $j$ -th eigenmodes, i.e.,  $\alpha_j \equiv \sum_n \mathcal{Q}_n x_j(n)$  and  $\beta_j \equiv \sum_n \mathcal{Q}_n y_j(n) P_n^{ss}$ , the correlation function is expanded in the eigenspace, i.e.,

$$C_Q(t - \tau) = \sum_{j=2}^N \alpha_j \beta_j e^{-\lambda_j |t - \tau|}. \quad (20)$$

The contribution of the ground state ( $j = 1$ ) cancels that of the mean deviation, i.e.,  $\langle Q \rangle_{ss}^2$ . Note that the stationarity leads to  $C_Q(t - \tau) = C_Q(\tau - t)$ , which constrains

Eq. (20) through the absolute term  $|t - \tau|$ . This can be derived by The correlation function  $C_{\dot{Q}}(t - \tau)$  for the velocity observable  $\dot{Q}(t)$  can be obtained by the transformation

$$C_{\dot{Q}}(t - \tau) = \frac{\partial^2 C_Q(t - \tau)}{\partial \tau \partial t}.$$

In the Fourier space, we have

$$\tilde{C}_{\dot{Q}}(\omega) = \sum_{j=2}^N 2\alpha_j \beta_j \lambda_j \left[ 1 - \frac{1}{1 + (\omega/\lambda_j)^2} \right], \quad (21)$$

which is generally valid for a system in NESS.

### C. Response spectrum

The response spectrum can be obtained by studying the response of the system to a periodic perturbation. Consider  $h = h_0 \exp(i\omega t)$  with  $h_0$  a small amplitude and  $i$  the imaginary unit. Up to the first order of  $h_0$ , the transition rate matrix  $\tilde{M}$  is modified as

$$\tilde{M} = M + h_0 \partial_h \tilde{M} \exp(i\omega t) + O(h_0^2).$$

After a sufficiently long time, the system reaches a distribution with a periodic temporal component that has time-independent amplitude:

$$\tilde{P}_m = P_m^{ss} + h_0 P_m^{(1)} \exp(i\omega t) + O(h_0^2).$$

Here, with the stationary condition of the zeroth term, i.e.,  $\sum_m M_{nm} P_m^{ss} = 0$ , the new master equation  $d\tilde{P}_m/dt = \sum_n \tilde{M}_{mn} \tilde{P}_n$  determines the first order correction of the distribution

$$P^{(1)} = -\frac{1}{M - i\omega} \partial_h M P^{ss},$$

written in a Matrix form. By introducing

$$B_n \equiv \sum_m \partial_h \tilde{M}_{nm} P_m^{ss}, \quad (22)$$

and the weighted average of  $B$  in the  $j$ -th eigenmode, i.e.,  $\phi_j \equiv \sum_n B_n y_j(n)$ , the linear order variation is expressed as

$$P_n^{(1)} = \sum_{j=2}^N \frac{1}{\lambda_j + i\omega} \phi_j x_j(n).$$

The first mode disappears as one can check that  $\phi_1 = \sum_m B_m y_1(m) = \sum_m B_m = 0$ , because the ground state does not contain dynamic information. Finally, for a state-dependent observable  $Q(t) \equiv Q_{n_t}$ , its response spectrum is given by

$$\tilde{R}_Q(\omega) = \sum_n Q_n P_n^{(1)} = \sum_{j=2}^N \frac{\alpha_j \phi_j}{\lambda_j + i\omega}. \quad (23)$$

By using the transformation  $R_{\dot{Q}}(t) = dR_Q/dt$  or  $\tilde{R}_{\dot{Q}}(\omega) = i\omega \tilde{R}_Q(\omega)$ , we obtain the desired response spectrum  $\tilde{R}_{\dot{Q}}(\omega)$  for the velocity observable  $\dot{Q}(t)$ , i.e.,

$$\tilde{R}_{\dot{Q}}(\omega) = \sum_{j=2}^N \alpha_j \phi_j \left[ 1 - \frac{1 - i(\omega/\lambda_j)}{1 + (\omega/\lambda_j)^2} \right]. \quad (24)$$

For the perturbation form in Eq. (5), we have

$$B_n = \sum_m [w_m^n P_m^{ss} + w_n^m P_n^{ss}] (Q_n - Q_m) / 2T, \quad (25)$$

which gives the flux fluctuation of the conjugate variable  $Q_n$  at state  $n$ .

### D. Useful relations

We find that the coefficients always satisfy the following relation

$$\sum_{j=2}^N \alpha_j (\lambda_j \beta_j - T \phi_j) = 0, \quad (26)$$

which is called the *sum rule*. See Appendix B for the derivation. Combining Eq. (24) and Eq. (21), this relation leads to FRR in high frequency domain ( $\omega \gg \lambda_N$ ), i.e.,

$$\tilde{C}_{\dot{Q}}(\omega) = 2T \tilde{R}'_{\dot{Q}}(\omega). \quad (27)$$

This is consistent with our intuition that when the frequency is much higher than the characteristic rate of the system the correlation and response spectrum of the system only reflects the property of the thermal bath that is in equilibrium.

Combining Eq. (21) and Eq. (24), the FRR violation spectrum for a velocity observable  $\dot{Q}$  can be generally written as

$$\tilde{C}_{\dot{Q}}(\omega) - 2T \tilde{R}'_{\dot{Q}}(\omega) = 2 \sum_{j=2}^N \alpha_j \frac{T \phi_j - \beta_j \lambda_j}{1 + (\omega/\lambda_j)^2}. \quad (28)$$

Although the involved coefficients and eigenvalues in the rhs are probably complex numbers, the summation over all the eigenmodes guarantees a real violation spectrum, as shown in Appendix A. In the limit  $\omega \rightarrow 0$ , FRR is also valid since both the correlation and response becomes zero, as evident from Eq. (21) and Eq. (24). The integral of the FRR violation, denoted as  $\Delta_Q$ , is given by

$$\begin{aligned} \Delta_Q &= \int_{-\infty}^{\infty} \left( \tilde{C}_{\dot{Q}}(\omega) - 2T \tilde{R}'_{\dot{Q}}(\omega) \right) \frac{d\omega}{2\pi} \\ &= \sum_j \lambda_j \alpha_j (T \phi_j - \beta_j \lambda_j). \end{aligned} \quad (29)$$

The Harada-Sasa equality suggests that this quantity is related to the dissipation through the motion of  $Q$ .

We also find that the detailed balance  $w_n^m P_n^{eq} = w_m^n P_m^{eq}$  is equivalent to

$$\lambda_j \beta_j^{eq} = T \phi_j^{eq}, \quad (30)$$

which ensures that FRR be satisfied in all frequency domain. This relation is a general result independent of the perturbation form proposed in Eq. (5), as proved in Appendix C. Therefore, we may also talk about detailed balance in the eigenspace, and an eigenmode contributes to dissipation only when it violates the detailed balance, according to Eq. (29).

## V. MARKOV PROCESSES WITH TIME SCALE SEPARATION

Now, we consider Markov processes with time scale separation. We assume that the state space can be grouped into  $K$  different coarse-grained subspaces, denoted as  $p$  or  $q$ . A microscopic state  $k$  ( $l$ ) within coarse-grained  $p$  ( $q$ ) is denoted as  $p_k$  ( $q_l$ ), which serves as an alternative to our previous state notation  $n$  or  $m$ . We assume fast relaxation ( $\sim \tau_f$ ) within the subspace of a coarse-grained state and slow ‘hopping’ ( $\sim \tau_s$ ) to other subspaces associated with a different coarse-grained state. The competition of these two processes defines a dimensionless parameter  $\epsilon \equiv \tau_f/\tau_s$ . A typical example of this Markov system is illustrated in FIG. 2. Such an assumption implies that the transition rate matrix of the system can be decomposed as

$$M_{p_k q_l} = \frac{1}{\epsilon} \delta_{pq} M_{kl}^p + M_{p_k q_l}^{(0)}, \quad (31)$$

where  $M_{kl}^p$  ( $\sim 1$ ) describes a rescaled ‘internal’ Markov process within the same coarse-grained state  $p$ , and  $M_{p_k q_l}^{(0)}$  for jumps connecting different coarse-grained states. The transition rate from state  $k$  to  $l$  for  $M^p$  is denoted as  $w_k^l(p)$  ( $\sim 1$ ).

The goal of this section is to obtain Eq. (46), concerning the structure of the FRR violation spectrum for a general observable in this system, and also Eq. (53) for an observable that could only resolve different coarse-grained state, which is of particular interest. Below, we start by obtaining the eigenmodes of  $M$  by perturbation theory.

### A. Perturbation analysis for eigenmodes

We write the eigenmodes and the corresponding eigenvalue as Taylor series in  $\epsilon$ , i.e.,

$$\begin{aligned} x_j &= x_j^{(0)} + \epsilon x_j^{(1)} + O(\epsilon^2) \\ y_j &= y_j^{(0)} + \epsilon y_j^{(1)} + O(\epsilon^2) \\ \lambda_j &= \epsilon^{-1} \lambda_j^{(-1)} + \lambda_j^{(0)} + O(\epsilon^2), \end{aligned}$$

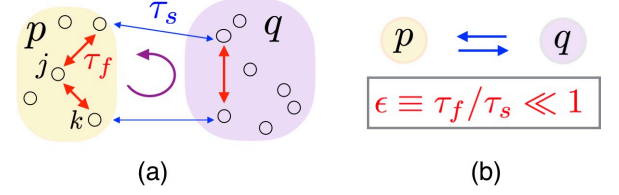


FIG. 2. (a) Our system with two time scales. One of the closed cycles that breaks time-reversal symmetry is indicated, which is responsible for hidden entropy production. (b) The corresponding effective dynamics.

and substitute these expansions into the eigenvalue equations Eq. (19). The leading order equations in  $\epsilon$  are given by

$$\sum_l M_{kl}^p x_j^{(0)}(p_l) = -\lambda_j^{(-1)} x_j^{(0)}(p_k) \quad (32a)$$

$$\sum_l y_j^{(0)}(p_l) M_{lk}^p = -\lambda_j^{(-1)} y_j^{(0)}(p_k). \quad (32b)$$

Therefore, the eigenmodes of the intra-block transition rate matrix  $\sum_p M^p$  are the same as the leading order eigenmodes of  $M$ . At  $\lambda_j^{(-1)} \neq 0$ , these eigenmodes are in general non-degenerate and decay quickly at the time scale  $\tau_f$ , thus called the fast modes.

The remaining  $K$  slow modes corresponding to  $\lambda_j^{(-1)} = 0$  are degenerate now, and the lift of this degeneracy is due to the inter-block transition, which requires the next order perturbation analysis, i.e.,

$$\begin{aligned} \sum_l M_{kl}^p x_j^{(1)}(p_l) + \sum_{q_l} M_{p_k q_l}^{(0)} x_j^{(0)}(q_l) &= -\lambda_j^{(0)} x_j^{(0)}(p_k) \\ \sum_l y_j^{(1)}(p_l) M_{lk}^p + \sum_{q_l} y_j^{(0)}(q_l) M_{q_l p_k}^{(0)} &= -\lambda_j^{(0)} y_j^{(0)}(p_k). \end{aligned}$$

Although the lhs (left hand side) depends on the unknown first order correction of the eigenmodes, we can eliminate these unknown terms by projecting the first equation on the left stationary mode of  $M^p$ , i.e., denoted as  $y_1^p = 1$ , and projecting the second equation on the right stationary mode of  $M^p$ , denoted as  $P(k|p)$  which satisfies the normalization  $\sum_k P(k|p) = 1$  and

$$\sum_k M_{lk}^p P(k|p) = 0. \quad (33)$$

We also introduce the following ansatz for these slow modes

$$x_j^{(0)}(p_k) = \hat{x}_j(p) P(k|p), \quad y_j^{(0)}(p_k) = \hat{y}_j(p), \quad (34)$$

which simply means that the eigenmodes are stationary under intra-block transition but have a modulation at the inter-block level. Then, we obtain a reduced eigenvalue

equations for these modulation amplitudes, i.e.,

$$\sum_q \widehat{M}_{pq} \widehat{x}_j(q) = -\lambda_j^{(0)} \widehat{x}_j(p) \quad (35a)$$

$$\sum_q \widehat{y}_j(q) \widehat{M}_{qp} = -\lambda_j^{(0)} \widehat{y}_j(p). \quad (35b)$$

Here, the emergent transition rate matrix on the coarse-grained state space is given by

$$\widehat{M}_{pq} \equiv \sum_{k,l} M_{pkql}^{(0)} P(l|q), \quad (36)$$

which is exactly due to a projection by the left and right stationary modes of the intra-block transition rate matrix. The projection procedure is effectively a coarse-graining over the microscopic states within the same block. Therefore, the effective matrix  $\widehat{M}$  removes the  $K$ -fold degeneracy of the slow modes, and determines its leading order behavior.

Now we summarize the non-degenerate leading order term of the eigenmodes of  $M$ . These eigenmodes are split into two classes in terms of its relaxation time scale: fast modes that relax at the time scale  $\sim \tau_f$  and slow ones at the time scale  $\sim \tau_s$ . This is illustrated in FIG. 2(c) for an illustrative example. For a fast mode, its leading order term is localized within a certain coarse-grained state  $p$ . We may denote the non-stationary eigenmodes of  $M^p$  as  $x_j^p$  and  $y_j^p$ , and the corresponding eigenvalue as  $\lambda_j^p$ . They satisfy the following orthogonal relations

$$\sum_k x_j^q(k) = 0, \quad \sum_k y_j^q(k) P(k|q) = 0. \quad (37)$$

The first relation can be understood from probability conservation, while the second one due to stationarity of the ground state. Then, this fast mode can be expressed as ( $j > K$ )

$$\lambda_j = \epsilon^{-1} \lambda_j^p + O(1) \quad (38a)$$

$$x_j(q_k) = \delta_{pq} x_j^p(k) + O(\epsilon) \quad (38b)$$

$$y_j(q_k) = \delta_{pq} y_j^p(k) + O(\epsilon). \quad (38c)$$

To express the slow modes, we introduce the eigenmodes of  $\widehat{M}$  as  $\widehat{x}_j$  and  $\widehat{y}_j$ , with corresponding eigenvalue  $\widehat{\lambda}_j$ . Then, the slow modes become ( $j \leq K$ )

$$\lambda_j = \widehat{\lambda}_j + O(\epsilon) \quad (39a)$$

$$x_j(p_k) = \widehat{x}_j(p) P(k|p) + O(\epsilon) \quad (39b)$$

$$y_j(p_k) = \widehat{y}_j(p) + O(\epsilon). \quad (39c)$$

In particular, the stationary distribution of the original system becomes

$$P_{pk}^{ss} = \widehat{P}_p^{ss} P(k|p) + O(\epsilon), \quad (40)$$

with  $\widehat{P}_p^{ss}$  the normalized stationary distribution for the coarse-grained Markov system.

The leading order results are sufficient for the following discussion. The first order correction is discussed in Appendix D. We note that the first order correction is in general very complicated. The correction of the fast modes involves only 1-step transition to all the other eigenmodes, while for the slow modes the correction also involves 2-step transition, i.e. first to a fast mode and then back to a different slow mode. The latter is generic in degenerate perturbation [40].

## B. FRR violation spectrum for fast observables

For a general conjugate variable  $Q_{p_k}^f$  that depends on microscopic states, its corresponding velocity observable  $\dot{Q}^f$  moves at a fast time scale  $\sim \tau_f$ , thus classified as a fast variable and emphasized by the superscript  $f$ . To obtain the property of its violation spectrum, we study the asymptotic behavior of the corresponding projection coefficients  $\alpha_j^f$ ,  $\beta_j^f$  and  $\phi_j^f$  with the help of the eigenmodes obtained in the previous section.

We first estimate their magnitude in the time scale separation limit  $\epsilon \rightarrow 0$ . Since the leading order terms for  $\alpha_j^f$  and  $\beta_j^f$  do not vanish for such a fast observable, we roughly obtain their magnitude as

$$\alpha_j^f \sim 1, \quad \beta_j^f \sim 1. \quad (41)$$

More delicate results up to first order correction can be obtained immediately by using Eq. (38) and Eq. (39). To obtain the magnitude of  $\phi_j^f \equiv \sum_{p_k} y_j(p_k) B_{p_k}$ , we need to understand  $B_{p_k}$  first. It can be expanded as [Eq. (25)]

$$B_{p_k} = \epsilon^{-1} B_k^p + B_{p_k}^{(1)}, \quad (42)$$

where the leading order term  $\epsilon^{-1} B_k^p \sim \epsilon^{-1}$  describes fast flux fluctuation within coarse-grained state  $p$ , as determined by  $M^p$ ; at the same time  $B_{p_k}^{(1)} \sim 1$  gives the slow flux fluctuation due to inter-block fluctuation, as determined by  $M^{(0)}$ . The projection of  $B_{p_k}$  on fast modes generally have a large value, i.e.,

$$\phi_j^f \sim \epsilon^{-1}, \quad j > K. \quad (43)$$

However, its projection on slow modes are greatly suppressed,

$$\phi_j^f \sim 1, \quad j \leq K, \quad (44)$$

because of the quasi uniformness of the slow modes in a given coarse-grained state, i.e.,  $y_j(p_k) \approx \widehat{y}_j(p)$ , and the conservation of flux fluctuation within each coarse-grained state, i.e.,  $\sum_k B_k^p = 0$ . The magnitude of these projection coefficients are listed in FIG. 3.

Since each fast mode has a counterpart from a certain internal transition rate matrix  $M^p$ , their projection coefficients also share this connection. For the Markov process described by  $M^p$ , we may also introduce the projection coefficients  $\alpha_j^q \equiv \sum_k x_j^q(k) Q_{qk}^f$ ,  $\beta_j^q \equiv$



$\sum_k y_j^q(k) \mathcal{Q}_{qk}^f P(k|q)$  and  $\phi_j^q \equiv \sum_k y_j^q(k) B(k|q)$ . Here,  $B(k|q)$  is the flux fluctuation for this subsystem, which is found to satisfy the relation  $B_k^q = \hat{P}_p^{ss} B(k|q) + O(\epsilon)$ . We find that these coefficients are connected to those of the fast mode by ( $j > K$ )

$$\alpha_j^f = \alpha_j^q + O(\epsilon) \quad (45a)$$

$$\beta_j^f = \hat{P}_q^{ss} \beta_j^q + O(\epsilon) \quad (45b)$$

$$\phi_j^f = \epsilon^{-1} \hat{P}_p^{ss} \phi_j^q + O(1). \quad (45c)$$

The FRR violation spectrum for this fast velocity observable  $\dot{Q}^f$  can be split into the contribution of each fast internal Markov processes [Eq. (28)] and a correction due to the slow transition across different coarse-grained state, i.e.,

$$\tilde{C}_{\dot{Q}^f} - 2T\tilde{R}'_{\dot{Q}^f} = \frac{2}{\epsilon} \sum_{j>K} \hat{P}_p^{ss} \left[ \alpha_j^q \frac{T\phi_j^q - \beta_j^q \lambda_j^q}{1 + (\epsilon\omega/\lambda_j^q)^2} \right] + V_f(\omega), \quad (46)$$

where the summation is first over all the non-stationary modes of  $M^p$ , and then over all the coarse-grained state  $p$ . The correction term is given by

$$V_f(\omega) = 2 \sum_{j \geq 2} \alpha_j^f \frac{T\phi_j^f - \beta_j^f \lambda_j^f}{1 + [\omega/\lambda_j^f]^2} - \frac{2}{\epsilon} \sum_{j>K} \hat{P}_p^{ss} \left[ \alpha_j^q \frac{T\phi_j^q - \beta_j^q \lambda_j^q}{1 + (\epsilon\omega/\lambda_j^q)^2} \right].$$

The two terms of the rhs have the same divergence of order  $\epsilon^{-1}$ , which cancels each other due to the mapping relation Eq. (45). Therefore,  $V_f(\omega)$  is of order 1 and is well-defined in the time scale separation limit  $\epsilon \rightarrow 0$ . Generally,  $V_f$  vanishes in the high frequency region  $\omega \gg \lambda_N$ , as expected from the sum rule. It also vanishes in the low frequency region for our setup with a finite state space. In the intermediate frequency region  $\tau_s^{-1} \ll \omega \ll \tau_f^{-1}$ , the contribute from slow modes is negligible and

$$V_f \simeq 2 \sum_{j>K} \alpha_j^f \left[ T\phi_j^f - \beta_j^f \lambda_j^f \right] \xrightarrow{\epsilon \rightarrow 0} \text{const.} \quad (47)$$

The finite limit is reached due to the mapping relations Eq. (45).

The FRR violation integral  $\Delta_{Q^f}$  also splits into two terms

$$\Delta_{Q^f} = \frac{1}{\epsilon} \sum_q \hat{P}_q^{ss} \Delta_{Q^f}^q + \Delta_{Q^f}^{(1)}, \quad (48)$$

where  $\Delta_{Q^f}^q$  is contributed by the FRR violation integral within mesoscopic state  $q$ , associated with  $M^q$ , and  $\Delta_{Q^f}^{(1)}$  is the correction term contributed by integrating over  $V_f(\omega)$ . Both  $\Delta_{Q^f}^q$  and  $\Delta_{Q^f}^{(1)}$  scale as  $\epsilon^{-1}$  since the violation plateau spans up to frequency  $1/\tau_f \sim \epsilon^{-1}$ . Therefore, the leading order term  $\frac{1}{\epsilon} \sum_q \hat{P}_q^{ss} \Delta_{Q^f}^q \sim \epsilon^{-2}$  and  $\Delta_{Q^f}^{(1)}$  is a negligible correction. Indeed, this leading order

		slow modes	fast modes
eigenvalue	$\lambda_j$	$\sim 1$	$\sim \epsilon^{-1}$
fast variable $\mathcal{Q}_{qk}^f$	$\alpha_j^f$	$\sim 1$	$\sim 1$
	$\beta_j^f$	$\sim 1$	$\sim 1$
	$\phi_j^f$	$\sim 1$	$\sim \epsilon^{-1}$
slow variable $\mathcal{Q}_p^s$	$\alpha_j^s$	$\sim 1$	$\sim \epsilon$
	$\beta_j^s$	$\sim 1$	$\sim \epsilon$
	$\phi_j^s$	$\sim 1$	$\sim 1$

FIG. 3. Overview of asymptotic behavior of the crucial parameters of correlation and response spectrum in this Markov system with time scale separation.

term implies a diverging dissipation rate of the system, which is not quite realistic. It is then natural to assume that detailed balance is satisfied within each coarse-grained state, i.e.,  $\phi_j^q = \beta_j^q \lambda_j^q$ , or equivalently  $\Delta_{Q^f}^q = 0$ . The FRR violation spectrum of this fast observable is then the same as  $V_f(\omega)$ , which is illustrated in FIG. 4(a).

### C. FRR violation spectrum for slow observables

Consider a special conjugate variable  $\mathcal{Q}_{p_k}^s = \mathcal{Q}_p^s$  that is uniform within the same coarse-grained state. It defines a velocity observable  $\dot{Q}^s(t) = \frac{d}{dt} \mathcal{Q}_{p_t}$  which is non-zero only when a slow transition to a neighboring coarse-grained state takes place, thus classified as a slow observable and emphasized by the superscript  $s$ . Below, we study the asymptotic behavior of its FRR violation spectrum for  $\epsilon \rightarrow 0$ . The main result has already been announced in our previous paper [24], especially in its supplemental material. Here, we provide more details.

First, we also estimate the magnitude of the corresponding projection coefficients  $\alpha_j^s$ ,  $\beta_j^s$ , and  $\phi_j^s$  in the limit  $\epsilon \rightarrow 0$ . The projection on slow modes gives a constant value, i.e., ( $j \leq K$ )

$$\alpha_j^s \sim 1, \quad \beta_j^s \sim 1, \quad (49)$$

which is not surprising according to their definition. However, the projection on fast modes is vanishingly small, i.e., ( $j > K$ )

$$\alpha_j^s \sim \epsilon, \quad \beta_j^s \sim \epsilon. \quad (50)$$

This is due to localization of the fast modes within certain coarse-grained states, and can be verified by using Eq. (37) and Eq. (38). The slow observable  $\dot{Q}^s$  is insensitive to flux fluctuation within coarse-grained states, which gives  $B_k^p = 0$  in Eq. (42). Therefore, we always have

$$\phi_j^s \sim 1, \quad (51)$$



whether it is projected on the slow or fast modes. The magnitude of these the projection coefficients are listed in FIG. 3.

Intuitively, the correlation and response of such a slow observable is pretty well described also on the coarse-grained level in terms of the effective Markov process  $\widehat{M}$ , which involves another set of projection coefficients  $\widehat{\alpha}_j^s \equiv \sum_p \widehat{x}_j(p) \mathcal{Q}_p^s$ ,  $\widehat{\beta}_j^s \equiv \sum_p \widehat{y}_j(p) \mathcal{Q}_p^s \widehat{P}_p^{ss}$ , and  $\widehat{\phi}_j^s \equiv \sum_p \widehat{y}_j(p) \widehat{B}_p$ . Here,  $\widehat{B}_p$  is the flux fluctuation defined for  $\widehat{M}$ , in the same spirit as  $B_n$  for  $M$  in Eq. (25), which satisfies  $\widehat{B}_p = \sum_k B_{pk}^{(1)} + O(\epsilon)$  due to the stationary distribution  $P_{pk}^{ss} = \widehat{P}_p^{ss} P(k|p) + O(\epsilon)$ . According to the mapping relations for slow modes in Eq. (39), the descriptions at the two levels are related by ( $j \leq K$ )

$$\alpha_j^s = \widehat{\alpha}_j^s + O(\epsilon) \quad (52a)$$

$$\beta_j^s = \widehat{\beta}_j^s + O(\epsilon) \quad (52b)$$

$$\phi_j^s = \widehat{\phi}_j^s + O(\epsilon), \quad (52c)$$

which are exactly the same in terms of the leading order of the slow modes. This justifies the validity of coarse-graining if we are interested in this slow observable.

The FRR violation spectrum of this slow observable can be split into the contribution from this coarse-grained description [Eq. (28)] and a correction from the underlying fast processes, i.e.,

$$\tilde{C}_{\dot{Q}^s} - 2T\tilde{R}'_{\dot{Q}^s} = 2 \sum_{j=2}^K \widehat{\alpha}_j^s \frac{T\widehat{\phi}_j^s - \widehat{\beta}_j^s \widehat{\lambda}_j}{1 + (\omega/\widehat{\lambda}_j)^2} + \epsilon V_s(\omega). \quad (53)$$

The correction term  $\epsilon V_s(\omega)$  is given by

$$\epsilon V_s(\omega) = 2 \sum_{j \geq 2} \alpha_j^s \frac{T\phi_j^s - \beta_j^s \lambda_j}{1 + (\omega/\lambda_j)^2} - 2 \sum_{j=2}^K \widehat{\alpha}_j^s \frac{T\widehat{\phi}_j^s - \widehat{\beta}_j^s \widehat{\lambda}_j}{1 + (\omega/\widehat{\lambda}_j)^2}.$$

With the mapping relations [Eq. (52)] for the slow modes and the magnitude estimation for the fast modes [FIG. 3], we find that  $\epsilon V_s(\omega)$  is of order  $\epsilon$ . Similar to  $V_f(\omega)$ ,  $V_s(\omega)$  also vanishes for both  $\omega \gg \lambda_N$  and  $\omega \ll \lambda_2$ . In the intermediate frequency region  $\tau_s^{-1} \ll \omega \ll \tau_f^{-1}$ , only the fast modes contributes, i.e.,

$$V_s \simeq \frac{2}{\epsilon} \sum_{j > K} \alpha_j^s (T\phi_j^s - \lambda_j \beta_j^s) \xrightarrow{\epsilon \rightarrow 0} \text{const}, \quad (54)$$

where the limit is obtained by using the magnitude estimation for the fast modes [FIG. 3]. Alternatively, we may express this plateau in terms of the slow modes by using the sum rule [Eq. (26)], i.e.,

$$V_s \simeq \frac{2}{\epsilon} \sum_{j=2}^K \alpha_j^s (\lambda_j \beta_j^s - T\phi_j^s) \xrightarrow{\epsilon \rightarrow 0} \text{const}. \quad (55)$$

The mapping relations [Eq. (52)] and the sum rule for the coarse-grained system, i.e.,  $\sum_{j=2}^K \widehat{\alpha}_j^s (\widehat{\lambda}_j \widehat{\beta}_j^s - T\widehat{\phi}_j^s) = 0$ ,

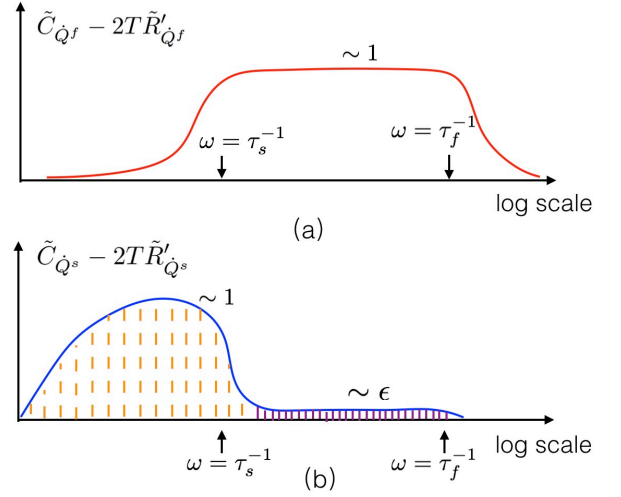


FIG. 4. (a) Illustration of FRR violation spectrum for a fast observable, assuming that detailed balance is satisfied within each coarse-grained state. This is equivalent to an illustration of  $V_f(\omega)$ , which is related to hidden entropy production. (b) Illustration of FRR violation spectrum for a slow observable. The violation in the low frequency region, shaded by orange dash line, is due to the broken of detailed balance at the level of effective dynamics, and the small violation at the intermediate frequency region is related to hidden entropy production.

guarantees this finite limit. See FIG. 4(b) for an illustration of  $\tilde{C}_{\dot{Q}^s} - 2T\tilde{R}'_{\dot{Q}^s}$ .

The FRR violation integral splits into two terms

$$\Delta_{Q^s} = \widehat{\Delta}_{Q^s} + \Delta_{Q^s}^{(1)}, \quad (56)$$

where the leading term  $\widehat{\Delta}_{Q^s}$  comes from the effective system  $\widehat{M}$  and  $\Delta_{Q^s}^{(1)}$  is the correction term from the integral of  $\epsilon V_s(\omega)$ . Although  $\epsilon V_s(\omega)$  is small, it extends to the high frequency cutoff  $1/\tau_f \sim \epsilon^{-1}$ , which makes the integral  $\Delta_{Q^s}^{(1)} \sim 1$ , comparable to the leading term  $\widehat{\Delta}_{Q^s}$ .

## VI. DISCUSSION

### A. Connection between the FRR violation plateau and hidden entropy production

It is shown by Esposito that steady state entropy production

$$\sigma \equiv \sum_{p_k, q_l} P_{pk}^{ss} w_{pk}^{ql} \ln \frac{P_{pk}^{ss} w_{pk}^{ql}}{P_{ql}^{ss} w_{ql}^{pk}} \quad (57)$$

for a Markov system with time scale separation ( $\epsilon \rightarrow 0$ ) can be splitted into three parts [22]: contribution from the coarse-grained dynamics

$$\sigma_1 = \sum_{p,q} \widehat{w}_p^q \widehat{P}_p^{ss} \ln \frac{\widehat{w}_p^q \widehat{P}_p^{ss}}{\widehat{w}_q^p \widehat{P}_q^{ss}}; \quad (58)$$

that from the microscopic transition within the same coarse-grained state

$$\sigma_2 = \frac{1}{\epsilon} \sum_p \hat{P}_p^{ss} \sum_{k,l} P(k|p) w_k^l(p) \ln \frac{P(k|p) w_k^l(p)}{P(l|p) w_l^k(p)}; \quad (59)$$

and that from the coupling between fast and slow transitions

$$\begin{aligned} \sigma_3 &\equiv \sigma - \sigma_1 - \sigma_2 \\ &= \sum_{p,q} \hat{w}_p^q \hat{P}_p^{ss} \sum_{k,l} f_{p_k}^{q_l} P(k|p) \ln \frac{f_{p_k}^{q_l} P(k|p)}{f_{q_l}^{p_k} P(l|q)}, \end{aligned} \quad (60)$$

where  $f_{p_k}^{q_l} \equiv w_{p_k}^{q_l} / \hat{w}_p^q$  is the conditional transition rate between microscopic state  $p_k$  and  $q_l$  given that transition between coarse-grained state  $p$  and  $q$  is already observed. All these three contributions are non-negative. Furthermore,  $\sigma_2$ , if exists, would diverge in the limit of timescale separation  $\epsilon \rightarrow 0$ .

In the steady state, the total entropy production rate of the system is equivalent to the total heat dissipation rate

$$J_{tot} \equiv T \sum_{p_k, q_l} P_{p_k}^{ss} w_{p_k}^{q_l} \ln \frac{w_{p_k}^{q_l}}{w_{q_l}^{p_k}}. \quad (61)$$

The reason is that  $\sigma - J_{tot}$  gives the change rate of the system entropy, i.e.,

$$\begin{aligned} \sigma - \frac{1}{T} J_{tot} &= \sum_{p_k, q_l} P_{p_k}^{ss} w_{p_k}^{q_l} \ln \frac{P_{p_k}^{ss}}{P_{q_l}^{ss}} \\ &= \sum_{p_k, q_l} [P_{p_k}^{ss} w_{p_k}^{q_l} - P_{q_l}^{ss} w_{q_l}^{p_k}] \ln P_{p_k}^{ss} \end{aligned}$$

which vanishes in the steady state because  $\sum_{q_l} [P_{p_k}^{ss} w_{p_k}^{q_l} - P_{q_l}^{ss} w_{q_l}^{p_k}] = 0$ . For the coarse-grained system, its total heat dissipation rate is given by

$$\hat{J}_{tot} = T \sum_{p,q} \hat{w}_p^q \hat{P}_p^{ss} \ln \frac{\hat{w}_p^q}{\hat{w}_q^p}. \quad (62)$$

For a coarse-grained state  $p$  that is assumed to be isolated from other coarse-grained states, its heat dissipation rate is given by

$$J_{tot}^p = \frac{1}{\epsilon} \sum_{k,l} P(k|p) w_k^l(p) \ln \frac{w_k^l(p)}{w_l^k(p)}. \quad (63)$$

Similar as  $\sigma = J_{tot}/T$ , we have

$$\sigma_1 = \frac{1}{T} \hat{J}_{tot}, \quad \sigma_2 = \frac{1}{T} \sum_p \hat{P}_p^{ss} J_{tot}^p. \quad (64)$$

Therefore,

$$\sigma_3 = \frac{1}{T} \left[ J_{tot} - \hat{J}_{tot} - \sum_p \hat{P}_p^{ss} J_{tot}^p \right]. \quad (65)$$

The above relations provide a chance to evaluate  $\sigma_1, \sigma_2$  and  $\sigma_3$  by quantifying the dissipation rates through FRR violation spectrum. However, a proper observable usually only captures part of the dissipation rate, as shown in Eq. (18). To capture the total dissipation rate requires very careful design of a set of “orthogonal” observables, with each taking care of dissipation rate from a subset of transitions in the network that are both complementary and non-overlapping. This also requires the application of the generalized Harada-Sasa equality, which is only approximately true for Markov jumping systems with proper transition rates as discussed before. Assuming that all these obstacles can be overcome, we can use the FRR violation integral for the effective dynamics, i.e.,  $\hat{\Delta}_{Q^s}$ , to evaluate  $\hat{J}_{tot}$ ,  $\Delta_{Q^f}^p$  for the fast dynamics within a coarse-grained state  $p$  to estimate  $J_{tot}^p$ , and the correction terms from the violation plateau, i.e.,  $\Delta_{Q^s}^{(1)}$  and  $\Delta_{Q^f}^{(1)}$  will contain information about  $\sigma_3$ .

Now, we assume that detailed balance is satisfied within each coarse-grained state, which implies that  $\sigma_2 = 0$  and  $J_{tot}^p = 0$ . In this case, the potential FRR violation for slow and fast observables are illustrated in FIG. 4. In FIG. 4(b) for the slow observables, one can clearly distinguish the orange area that contributes to  $\sigma_1$  and the purple area that contributes to  $\sigma_3$ . The plateau in FIG. 4(a) only contributes to  $\sigma_3$ . Therefore,  $\sigma_1$  comes from the FRR violation in the low frequency region  $\omega \sim \tau_s^{-1}$ , and  $\sigma_3$  comes from the FRR violation in the high frequency region  $\omega \sim \tau_f^{-1}$  (note that this plateau is plotted in a log scale and its integral is actually dominated by the region  $\omega \sim \tau_f^{-1}$ ). Then, it is possible to quantify  $\sigma_1$  and  $\sigma_3$  only from measurable quantities of the original system. This will be illustrated later through an Markov jumping example.

Below, under the assumption that  $\sigma_2 = 0$ , we connect  $\sigma_1$  and  $\sigma_3$  with FRR violation spectrum for the  $N_0$ -component Langevin equation (1). For this system, not only the Harada-Sasa equality is applicable, the complete set of orthogonal observables are also well-defined, which is  $\hat{x}_j$  for  $j = 1, 2, \dots, N_0$ . We assume that the variables are indexed in such a way that  $\gamma_j \sim 1$  for  $j \leq K_0$  and  $\gamma_j \sim \epsilon$  for  $j > K_0$ . This means that  $x_j$  is a slow variable for  $j \leq K_0$ , with  $K_0$  the number of slow variables in this system. We also assume that there is no net drift, i.e.,  $\langle x_j \rangle = 0$ . This multi-component Langevin system can be described by our general Markov system with time scale separation, where the coarse-grained state would be any configuration specified by the slow variables, i.e.,  $\vec{x}^s \equiv (x_1, x_2, \dots, x_{K_0})$ , and the microscopic state within  $\vec{x}^s$  would be any configuration specified by all the fast variables. With these assumptions, we have

$$\sigma_1 = \frac{1}{T} \hat{J}_{tot} = \frac{1}{T} \sum_{j=1}^{K_0} \gamma_j \hat{\Delta}_{x_j}, \quad (66a)$$

$$\sigma_3 = \frac{1}{T} \sum_{j=1}^{K_0} \gamma_j \Delta_{x_j}^{(1)} + \frac{1}{T} \sum_{j=K_0+1}^{N_0} \gamma_j \Delta_{x_j}. \quad (66b)$$

Here,  $\hat{\Delta}_{x_j}$  and  $\Delta_{x_j}^{(1)}$ , with  $j = 1, 2, \dots, K_0$ , can be evaluated by the integral over the low and high frequency region of the FRR violation spectrum for this slow observable  $\dot{x}_j$ , respectively.

For many physical systems, time scale separation usually implies that not only  $\sigma_2 = 0$  but also  $\sigma_3 = 0$ , i.e., no hidden entropy production at all. An interesting example is the potential switching model for molecular motors [41], where chemical transition is fast and displacement is slow. In this scenario, the total dissipation rate could be extracted by only studying the FRR violation of the slow observables, as long as the slow transitions satisfy our assumptions that lead to Eq. (18).

### B. Effective temperature for a fast observable

Here, we define effective temperature by naively assuming the FRR for all frequency:

$$T_{eff}(\omega; \dot{Q}) \equiv \frac{\tilde{C}_{\dot{Q}}(\omega)}{2\tilde{R}'_{\dot{Q}}(\omega)}. \quad (67)$$

Note that this definition should be modified accordingly to apply for a displacement observable  $Q(t)$ , so as to give the same result. In general,  $T_{eff}$  depends both on frequency and the observable considered. It converges to the bath temperature in the high frequency region  $\omega \gg \lambda_N$ .

This concept becomes popular after Cugliandolo *et al.* apply it to glassy systems in [29], where they argue that the inverse effective temperature actually controls the direction of heat flow. For example,  $1/T_{eff}(\omega, \dot{Q}) < 1/T$  implies that heat flows from this degree of freedom  $Q(t)$  to the bath. Although  $T_{eff}$  share this desirable property with the temperature of an equilibrium system, its value is in general sensitive to the choice of observable [42]. Below, we discuss the property of this definition for our system. Its physical meaning will be treated in elsewhere.

We assume that  $\sigma_2 = 0$ , which implies that  $T\phi_j^q = \beta_j^q \lambda^q$  for transitions within the same coarse-grained state. Then, the effective temperature of a general fast observable is given by

$$T_{eff}(\omega, \dot{Q}^f) = T + \frac{V_f(\omega)}{2\tilde{R}'_{\dot{Q}^f}(\omega)}. \quad (68)$$

At the frequency region  $\omega \gg \tau_f^{-1/2}$ , the response spectrum  $\tilde{R}'_{\dot{Q}^f}(\omega) \sim \epsilon^{-1}$ , dominated by the fast modes, while at the frequency  $\omega \ll \tau_f^{-1/2}$  we have  $\tilde{R}'_{\dot{Q}^f}(\omega) \sim 1$ , dominated by the slow modes. Besides,  $V_f(\omega)$  changes from zero to a finite value around the frequency  $\tau_s^{-1}$ . Combining these, we find that this effective temperature takes constant value in three frequency regions:

$$T_{eff}(\omega, \dot{Q}^f) = \begin{cases} T + T_1^f, & \omega \ll \tau_s^{-1} \\ T + T_2^f, & \tau_s^{-1} \ll \omega \ll \tau_f^{-1/2} \\ T, & \omega \gg \tau_f^{-1/2}. \end{cases} \quad (69)$$

Here,  $T_1^f$  and  $T_2^f$  are two different numbers that satisfy, up to leading order,

$$T_1^f = \frac{\sum_{j=2}^K \alpha_j^f (\beta_j^f \lambda_j - T\phi_j^f) / \lambda_j^2}{\sum_{j=2}^K \alpha_j^f \phi_j^f / \lambda_j^2} + O(\epsilon), \quad (70a)$$

$$T_2^f = \frac{\sum_{j=2}^K \alpha_j^f (\beta_j^f \lambda_j - T\phi_j^f)}{\sum_{j=2}^K \alpha_j^f \phi_j^f} + O(\epsilon). \quad (70b)$$

Note that  $T_1^f$  has a similar structure as  $T_2^f$ , except for a weighting factor  $1/\lambda_j^2$  that is ordered in a descending way, i.e.,  $1/\lambda_2^2 \geq 1/\lambda_3^2 \geq \dots$ . Both  $T_1^f$  and  $T_2^f$  pick up an eigenmode that contributes dominantly. However,  $T_1^f$  favors a slower eigenmode due to the weighting factor. Only a eigenmode that violates detailed balance can contribute, i.e.,  $\beta_j^f \lambda_j \neq T\phi_j^f$ .

Through quite a few non-trivial examples, we find that

$$T_1^f \approx T_2^f. \quad (71)$$

This implies that both  $T_1^f$  and  $T_2^f$  pick up the slowest eigenmode. In other words, the slowest eigenmode ( $j = 2$ ) contributes more significantly to the violation of detailed balance than other slow modes, i.e., for  $2 < j \leq K$ ,

$$|\alpha_2^f (\beta_2^f \lambda_2 - T\phi_2^f)| > |\alpha_j^f (\beta_j^f \lambda_j - T\phi_j^f)|. \quad (72)$$

This agrees with our intuition that slower modes takes longer time to relax to equilibrium and thus breaks detailed balance more easily when driven out of equilibrium. Eq. (71) implies that this two-timescale system has only two distinct temperatures at the large and slow timescale respectively, which are independent of the time scale separation index  $\epsilon$ .

### C. Effective temperature for a slow observable

The effect of hidden fast processes on a slow observable can be captured by a renormalized (effective) rate matrix Eq.(36) and a fast colored noise with small correlation time  $\sim \tau_f$ . We assume that the effective dynamics  $\hat{M}$  in the slow time scale  $\tau_s$  satisfies detailed balance, and seek to capture the noise effect by an effective temperature.

The effective temperature for a slow observable  $Q^s$  consists of a constant bath temperature  $T$  and a frequency-dependent component from the colored noise:

$$T_{eff}(\omega, \dot{Q}^s) = T + \epsilon \frac{V_s(\omega)}{2\tilde{R}'_{\dot{Q}^s}(\omega)}. \quad (73)$$

In general,  $\tilde{R}'_{\dot{Q}^s}(\omega) \sim 1$  at all frequency, dominated by the slow modes. Besides,  $V_s(\omega) \sim 1$  and has two cut-offs at the low frequency  $\tau_s^{-1}$  and high frequency  $\tau_f^{-1}$ ,

respectively. Therefore, we find that the effective temperature becomes constant in the low, intermediate, and high frequency region:

$$T_{eff}(\omega; \dot{Q}^s) = \begin{cases} T + \epsilon T_1^s, & \omega \ll \tau_s^{-1} \\ T + \epsilon T_2^s, & \tau_s^{-1} \ll \omega \ll \tau_f^{-1} \\ T, & \omega \gg \tau_f^{-1}. \end{cases} \quad (74)$$

Here,  $T_1^s$  ( $T_2^s$ ) has a similar structure as  $T_1^f$  ( $T_2^f$ ), i.e.,

$$T_1^s = \frac{1}{\epsilon} \frac{\sum_{j=2}^K \alpha_j^s (\beta_j^s \lambda_j - T \phi_j^s) / \lambda_j^2}{\sum_{j=2}^K \alpha_j^s \phi_j^s / \lambda_j^2} + O(\epsilon^2), \quad (75)$$

$$T_2^s = \frac{1}{\epsilon} \frac{\sum_{j=2}^K \alpha_j^s (\beta_j^s \lambda_j - T \phi_j^s)}{\sum_{j=2}^K \alpha_j^s \phi_j^s} + O(\epsilon^2), \quad (76)$$

where  $\beta_j^s \lambda_j - T \phi_j^s \sim \epsilon$  due to our assumption that the system reach equilibrium effectively in the large time scale region, which ensures that  $T_1^s$  and  $T_2^s$  are of order 1. The slow dynamics is frozen at the intermediate frequency region  $\tau_s^{-1} \ll \omega \ll \tau_f^{-1}$ , while it evolves to a stationary distribution at the time scale much larger than  $\tau_s$ . Since the strength of this colored noise generally depends on the value of the slow variable, it is generally renormalized into different noise strength at the intermediate and slow frequency region, respectively.  $T_1^s$  or  $T_2^s$  measures the strength of this noise, and thus strength of the driving from the hidden fast processes. In the limit that  $\epsilon \rightarrow 0$ , the effective temperature converges to the bath temperature.

We also find that

$$T_1^s \approx T_2^s. \quad (77)$$

The underlying mechanism is similar to that of Eq. (71). Eq. (77) implies that we can roughly parameterize the active noise effect by an constant amplitude in the whole frequency region  $\omega \ll \tau_f^{-1}$ . In other words, the system has the same temperature with the bath at the time scale smaller than  $\tau_f$ , and a slightly different temperature  $T + \epsilon T_1^s$  at the time scale larger than  $\tau_f$ .

## VII. EXAMPLE: SENSORY ADAPTATION NETWORK

Here, we study the Markov network illustrated in FIG. 5, which describes sensory adaptation of the membrane receptor in E.coli [5, 18, 43]. This network contains two degrees of freedom:  $a (= 0, 1)$  for the activity of this protein and  $m (= 0, 1, \dots, m_0)$  for its methylation level. Here,  $a$  changes relatively fast on the time scale  $\tau_f$  while  $m$  changes on the slow time scale  $\tau_s$ . For a fixed  $m$ , the activity reaches a local equilibrium distribution  $P(1|m)/P(0|m) = \exp(-\Delta S(m))$ , where  $\Delta S(m)$  is a linear function given by  $\Delta S(m) = e_0(m_1 - m)$ . In general, the effect of extracellular ligand binding in NESS is

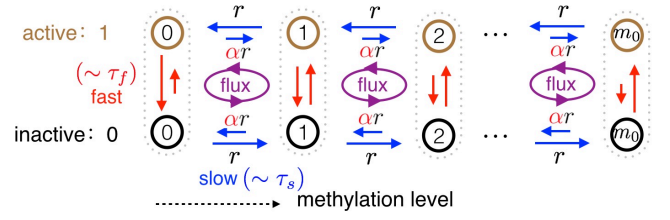


FIG. 5. Sensory adaptation network for a single membrane receptor in E.coli. The activity of this receptor  $a = 1$  in the active state and 0 when inactive. Its methylation level  $m$  ranges from 0 to  $m_0$ . Actually,  $m_0 = 4$  for this receptor in E.coli. Here  $\alpha \ll 1$  and  $\tau_s \gg \tau_f$  to achieve adaptation. Each methylation level is regarded as a coarse-grained state that contains two microscopic states with different activity.

captured by a shift of  $m_1$ . We may assume that the activation rate  $w_0(m)$  and inactivation rate  $w_1(m)$  satisfy

$$w_0(m) = \frac{1}{\tau_f} \exp\left(-\frac{\Delta S(m)}{2T}\right), \quad (78a)$$

$$w_1(m) = \frac{1}{\tau_f} \exp\left(\frac{\Delta S(m)}{2T}\right). \quad (78b)$$

We assume that in the inactive (active) conformation the methylation (demethylation) rate is  $r$ , while the reverse transition rate is attenuated by a small factor  $\alpha$ , as illustrated in FIG. 5. The time scale of methylation events is given by  $\tau_s = 1/r$ .  $\alpha \ll r$  is required for high sensory adaptation accuracy in E.coli. The time scale separation of this system is again captured by  $\epsilon \equiv \tau_f/\tau_s$ , which should be small to achieve adaptation. For more details of this model, please refer to [18].

First, we study the fast observable  $\dot{a}$ , the change rate of the activity. To obtain its response, we apply a perturbation  $h$  that increase the activation rate  $w_0$  by a factor  $\exp(h/2T)$ , but decrease the inactivation rate  $w_1$  by a factor  $\exp(-h/2T)$ . Such a perturbation equivalently modifies  $\Delta S(m) \rightarrow \Delta S(m) - h$ , which can be realized by changing the extracellular ligand concentration slightly. According to the proposed perturbation form Eq. (5), the activity  $a$  is the corresponding conjugate field.

FIG. 6(a) shows the numerically exact velocity correlation spectrum  $\tilde{C}_{\dot{a}}$  and the real part of the response spectrum  $\tilde{R}'_{\dot{a}}$  at various  $\epsilon$ . We can see that the FRR is approximately satisfied in the high frequency region  $\omega \sim \tau_f^{-1}$ , but violated in the low frequency region  $\omega \sim \tau_s^{-1}$ . In particular, the response spectrum  $\tilde{R}'_{\dot{a}}$  becomes negative in this low frequency region, while the correlation spectrum  $\tilde{C}_{\dot{a}}$  remains positive, resulting in a negative effective temperature in this low frequency region. Still, the inverse effective temperature  $1/T_{eff}(\omega, \dot{a}) \leq 1/T$  at all the frequency, and therefore the heat also flows from this degree of freedom  $a$  to the bath even at this low frequency region. This is illustrated in FIG. 6(b), where the effective temperature becomes a negative constant for low frequency region  $\omega \ll \tau_f^{-1/2}$  and reaches bath tem-

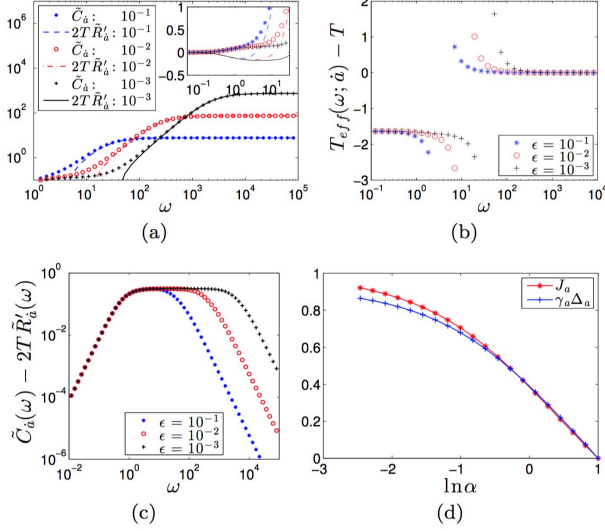


FIG. 6. (a) Correlation and response spectrum for the fast observable  $\dot{a}$  at various  $\epsilon$ . The inset shows discrepancy of the two spectrum at the low frequency region. (b) The effective temperature of this observable. (c) The corresponding FRR violation spectrum. (d) The dissipation rate  $J_a$  due to change of only activity and the estimate  $\gamma_a \Delta_a$  by assuming Harada-Sasa equality, with  $\gamma_a = 1.4\tau_f T$ . Parameter:  $r = 1$  (thus,  $\epsilon = \tau_f$ ),  $\alpha = 0.1$ ,  $e_0 = 2$ ,  $T = 1$ ,  $m_0 = 4$ , and  $m_1 = m_0/2$ . We vary  $\tau_f$  to change  $\epsilon$ . For (d), we fix  $\tau_f = 0.1$  and change  $\alpha$  instead.

perature for high frequency region  $\omega \gg \tau_f^{-1/2}$ . FIG. 6(c) shows that the corresponding FRR violation spectrum has a plateau in the broad intermediate frequency region  $\tau_s^{-1} \ll \omega \ll \tau_f^{-1}$  with an  $\epsilon$ -independent amplitude, which agrees with our general analysis. Although this plateau is of order 1, it is much smaller compared with the correlation or response spectrum in the high frequency region, which is of order  $\epsilon^{-1}$ . Indeed, the frequency-dependence of the effective temperature suggests that the FRR violation is much easier to be detected in the low frequency region.

Next, we consider the slow observable  $\dot{m}$ , the change rate of the methylation level. To obtain its response, we use a perturbation  $h$  that increases all the methylation rate (i.e.,  $r$  for inactive state and  $\alpha r$  for active state) by a factor  $\exp(h/2T)$ , but decreases all the demethylation rate (i.e.,  $\alpha r$  for inactive state and  $r$  for active state) by a factor  $\exp(-h/2T)$ . The conjugate field here is the methylation level  $m$ .

The numerically exact correlation and response spectrum are shown in FIG. 7(a). The violation of FRR appears mainly in the intermediate frequency range and this violation tends to vanish in the limit  $\epsilon \rightarrow 0$ , implying an equilibrium-like dynamics in the time scale separation limit. The non-equilibrium effect of the hidden fast variable can be captured by the extra effective temperature  $T_{eff} - T$ , as shown in FIG. 7(b), which is almost constant in the region  $\omega \ll \tau_f^{-1}$  and has a vanishingly small

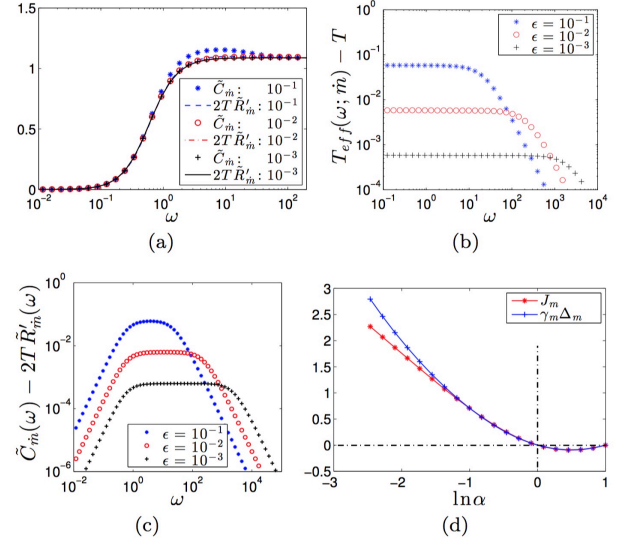


FIG. 7. The analysis for the observable  $\dot{m}$ , similar to FIG. 6. Here,  $\gamma_m = T(r\sqrt{\alpha})^{-1}$ . The parameters are the same as those in FIG. 6.

amplitude that scales linearly with  $\epsilon$ . The positivity of this extra effective temperature gives rise to a small heat flow from the degree of freedom  $m$  to the bath at all frequency. The FRR violation spectrum serves as another measure of the non-equilibrium effect of the hidden fast processes, as is shown in FIG. 7(c). It has a plateau in the frequency region  $\tau_s^{-1} \ll \omega \ll \tau_f^{-1}$  with also a small amplitude of order  $\epsilon$ . This behavior confirms our general analysis. This FRR violation can be quite difficult to detected experimentally.

The total dissipation rate of a Markov system is given by

$$J_{tot} = \frac{1}{2} \sum_{n,n'} (P_n^{ss} w_n^{n'} - P_{n'}^{ss} w_{n'}^n) \ln \frac{w_n^{n'}}{w_{n'}^n}, \quad (79)$$

which is always non-negative in the stationary state. In our bipartite network, the total dissipation can be decomposed into dissipation for change of activity, denoted as  $J_a$ , and for change of methylation level, denoted as  $J_m$ . For example,

$$J_a = \sum_m [P_{1,m}^{ss} w_1(m) - P_{0,m}^{ss} w_0(m)] \ln \frac{w_1(m)}{w_0(m)}. \quad (80)$$

Then, the total entropy production  $\sigma$  in our system satisfies

$$\sigma = \sigma_3 = \frac{1}{T} (J_a + J_m), \quad (81)$$

where the first equality holds because our system has an equilibrium effective dynamics for the slow variable  $m$ , i.e.,  $\sigma_1 = 0$ , and the fast dynamics for each given methylation level also satisfies detailed balance, i.e.,  $\sigma_2 = 0$ . The result of  $J_a$  and  $J_m$  is shown in FIG. 6(d) and



FIG. 7(d), respectively. Here, we fix  $\tau_f = 0.01$  and change  $\alpha$  to tune the dissipation rate of the system. Note that  $J_m > 0$  for  $\alpha < 1$  and  $J_m < 0$  for  $1 < \alpha < \exp(1)$ , which implies that  $\alpha = 1$  is a critical point of the system [18].

The generality of Harada-Sasa equality in Langevin systems motivates us to check the following relation between FRR violation and dissipation in this bipartite network:

$$J_a \approx \gamma_a \Delta_a, \quad J_m \approx \gamma_m \Delta_m, \quad (82)$$

where  $\Delta_a$  and  $\Delta_m$  are the FRR violation integral for the observable  $\hat{a}$  and  $\hat{m}$ , respectively. And,  $\gamma_a$  and  $\gamma_m$  are the corresponding effective friction coefficients, which is in general not well-defined in Markov systems. We may still derive  $\gamma_m = T(r\sqrt{\alpha})^{-1}$  by focusing on the region that  $\alpha$  is close to 1, where the methylation dynamics is almost diffusive and can be well-approximated by a Langevin equation. See Supplemental Material in [24] for details. However,  $\gamma_a$  cannot be derived by a similar method because we do not have a Langevin-analogy for this two-state network (with given  $m$ ). A naive way to overcome this difficulty is to define  $\gamma_a = J_a/\Delta_a$  at a particular set of parameters, and then use this  $\gamma_a$  to check whether  $J_a \approx \gamma_a \Delta_a$  holds in a more broad parameter region. In this way, we take  $\gamma_a = 1.4\tau_f T$ , determined from  $\tau_f = 0.1$  and  $\alpha = 1$ .

With these two effective friction coefficients, we numerically compare  $J_x$  and  $\gamma_x \Delta_x$  ( $x = a, m$ ) in FIG. 6(d) and FIG. 7(d), by fixing  $\tau_f$  and varying  $\alpha$ . Both agree very well in the region  $\exp(-1) \leq \alpha \leq \exp(1)$ , which is very non-trivial because the system displays qualitatively very different behavior for  $\alpha > 1$  (no adaptation with strong boundary effect from  $m = 0, m_0$ ) and for  $\alpha < 1$  (adaptation with negligible boundary effect from  $m = 0, m_0$ ) [18]. These relations also hold at various  $\epsilon$ , as shown in FIG. 8. The Harada-Sasa equality holds approximately for the observable  $\hat{m}$  because the methylation dynamics satisfies the assumptions underlying our generalized Harada-Sasa equality (18), in particular, a relatively diffusive transition rate. However, the validity of  $J_a \approx \gamma_a \Delta_a$  would be more difficult to understand, because this observable has only two states (i.e.,  $a = 0, 1$ ) and the two state dynamics is definitely not quite diffusive since  $\Delta S(m)$  varies broadly.

Below, we investigate in detail how the Harada-Sasa equality works for the observable  $\hat{a}$ . According to Eq. (8) and Eq. (10), we have

$$\Delta_a = \sum_m [P_{1,m}^{ss} w_1(m) - P_{0,m}^{ss} w_0(m)] \frac{w_1(m) - w_0(m)}{2}. \quad (83)$$

We define  $\Delta_a(m)$  to be the contribution of  $m$ -th methylation level to  $\Delta_a$ , and therefore  $\Delta_a = \sum_m \Delta_a(m)$ . Similarly, we define  $J_a(m)$  to be the contribution of  $m$ -th methylation level to  $J_a$  in Eq. (80), which satisfies  $J_a = \sum_m J_a(m)$ . FIG. 8(b) shows that  $J_a(m) \approx \gamma_a \Delta_a(m)$ , which is necessary to explain why  $J_a \approx \gamma_a \Delta_a$  works over

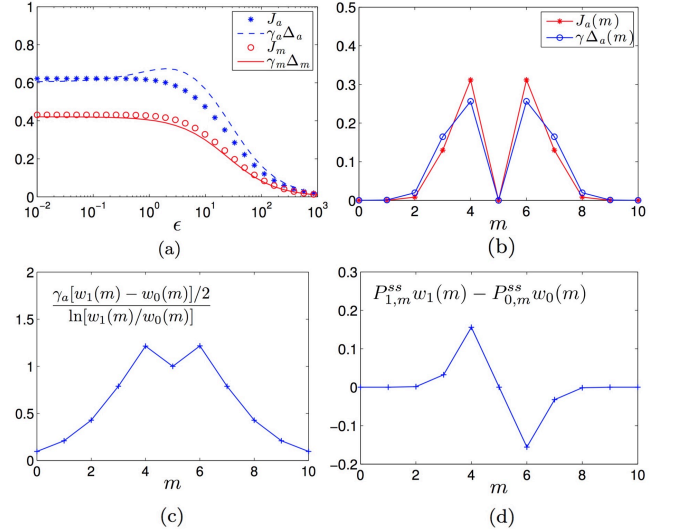


FIG. 8. (a) Test of generalized Harada-Sasa equality Eq. (82) at different  $\epsilon$ . Major violation happens only at  $\epsilon \approx 1$ . Here,  $\alpha = 0.5$  and  $\tau_f$  is varied to change  $\epsilon$ . Other parameters are the same as FIG. 6. (b) Comparison between  $J_a(m)$  and  $\gamma_a \Delta_a(m)$ . (c) Distribution of the ratio between  $\gamma_a(w_1(m) - w_0(m))/2$ , an element from  $\gamma_a \Delta_a(m)$ , and  $\ln[w_1(m)/w_0(m)]$ , an element in  $J_a(m)$ . (d) Activity flux distribution. In (b)(c)(d), the parameters are  $\alpha = 0.1$ ,  $m_0 = 10$  and  $\tau_f = 0.1$ . Other parameters are the same as FIG. 6.

such a broad parameter region. Further analysis reveals that  $\gamma_a(w_1(m) - w_0(m))/2 \approx \ln[w_1(m)/w_0(m)]$  holds only for a narrow methylation region where the dissipation  $|\Delta S(m)|$  is relatively small, as shown in FIG. 8(c). Fortunately, the activity flux  $[P_{1,m}^{ss} w_1(m) - P_{0,m}^{ss} w_0(m)]$  is significant only around the region where  $|\Delta S(m)|$  is also relatively small, as shown in FIG. 8(d). This explains why  $J_a(m) \approx \gamma_a \Delta_a(m)$  holds.

## VIII. CONCLUSION

Here, we have made a systematic analysis for a general Markov system with two time scales, focusing on the FRR violation spectrum. The two characteristic time scales divide the frequency region into three domains: the low, intermediate, and high frequency region. Even assuming that the fast processes satisfy detailed balance, the FRR violation for either a slow or a fast observable is characterized by a plateau in the intermediate frequency region. Generically, this plateau implies a finite hidden entropy production rate that results from coupling between slow and fast processes. This connection is formulated precisely for general Langevin systems of two time scales. A very interesting Markov jumping system motivated from sensory adaptation in E.coli also supports this connection. To quantify hidden entropy production from FRR violation spectrum, we need to properly choose a complete set of orthogonal observables that

capture all independent channels of dissipation, and measure the FRR violation plateau for each of them.

We have also studied a different measure of non-equilibrium dynamics: effective temperature. For a NESS only driven by the coupled motion between fast and slow processes, we find that the effective temperature for a fast observable approaches the room temperature in the high frequency region, while it significantly deviates from the room temperature in the low frequency region. This two-temperature two-time scale scenario is similar with glassy systems [29, 44]. However, the effective temperature for a slow observable approaches the bath temperature throughout all frequency region in the timescale separation limit, which is consistent with the emergence of an effective potential landscape. The tiny deviation of order  $\epsilon$  from the bath temperature appears at the low and intermediate frequency region, which is crucial to explain the finite dissipation rate of the slow variable. This extra deviation could be modeled by an extra fast noise, which would then capture the feature of a finite dissipation rate. The above results suggest that it is much easier to probe hidden entropy production by measuring the low frequency violation of the FRR for a fast observable.

The Harada-Sasa equality was originally derived only for Langevin dynamics with an infinite state space. Here, we also present a systematic discussion on the applicability of this equality to general Markov jumping systems. We find that the generalized Harada-Sasa equality not only requires a relatively *diffusive* transition along the direction of the observable, the prefactor of the transition rate, which quantifies its time scale, should also be homogeneous along this direction. Here, a transition is called to be diffusive when it produces only a small amount of entropy in the medium. These requirements allow to lump all the transitions in this direction together, and therefore to access their dissipation rate by only monitoring the stochastic evolution projected along this direction. Other details such as the system size are not rel-

evant. In some cases, although the relevant transitions are not always diffusive, the transitions that are more diffusive may take place much more frequently, which may again restore this equality, as supported by our sensory adaptation model. This generalized Harada-Sasa equality can be very useful to measure the total entropy production rate for a system with time scale separation, because the fast processes usually reach equilibrium and that the phenomenon of hidden entropy production is not so common in physical systems.

Our Markov system assumes a finite state space, which forbids a net drift of any observables. However, many interesting periodic systems are able to reach NESS and at the same time have a drifting motion, say molecular motors. For such systems, another key difference is that the correlation and response spectrum will not vanish at the low frequency limit, and consequently FRR violation at the low frequency limit is also possible. However, the other features identified in our current paper will remain unchanged. This discussion will be presented elsewhere.

In the near future, it would be more interesting to apply our general framework to interesting systems with hidden entropy production, say (possibly) inefficiently molecular motors [45], active cytoskeletal networks [27], and repulsive self-propelled particles [46], where gas and liquid phase coexist.

## ACKNOWLEDGEMENTS

The authors thank M. Esposito and Y. Nakayama for fruitful discussions and comments. The work was supported in part by the NSFC under Grant No. U1430237 and by the Research Grants Council of the Hong Kong Special Administrative Region (HKSAR) under Grant No. 12301514. It was also supported by KAKENHI (Nos. 25103002 and 26610115), and by the JSPS Core-to-Core program “Non-equilibrium dynamics of soft-matter and information”.

- 
- [1] K. Sekimoto, in *Lecture Notes in Physics, Berlin Springer Verlag*, Vol. 799 (2010).
  - [2] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
  - [3] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **71**, 4135 (1974).
  - [4] P. Sartori and S. Pigolotti, Phys. Rev. Lett. **110**, 188101 (2013).
  - [5] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, Nat. Phys. (2012).
  - [6] P. Mehta and D. J. Schwab, Proc. Natl. Acad. Sci. USA **109**, 17978 (2012).
  - [7] S. Bo, M. D. Giudice, and A. Celani, Journal of Statistical Mechanics: Theory and Experiment **2015**, P01014 (2015).
  - [8] Y. Cao, H. Wang, Q. Ouyang, and Y. Tu, Nat. Phys. **11**, 772 (2015).
  - [9] H. Noji, R. Yasuda, M. Yoshida, and K. Kinoshita, Nature **386**, 299 (1997).
  - [10] S. Toyabe, T. Watanabe-Nakayama, T. Okamoto, S. Kudo, and E. Muneyuki, Proc. Natl. Acad. Sci. U.S.A. **108**, 17951 (2011).
  - [11] S. Rahav and C. Jarzynski, J. Stat. Mech. Theor. Exp. **2007**, P09012 (2007).
  - [12] A. Puglisi, S. Pigolotti, L. Rondoni, and A. Vulpiani, J. Stat. Mech. Theor. Exp. **2010**, P05015 (2010).
  - [13] A. Celani, S. Bo, R. Eichhorn, and E. Aurell, Phys. Rev. Lett. **109**, 260603 (2012).
  - [14] S. Bo and A. Celani, J. Stat. Phys. **154**, 1325 (2014).
  - [15] Y. Nakayama and K. Kawaguchi, Phys. Rev. E **91**, 012115 (2015).
  - [16] H.-M. Chun and J. D. Noh, Phys. Rev. E **91**, 052128 (2015).



- [17] M. Esposito and J. M. R. Parrondo, Phys. Rev. E **91**, 052114 (2015).
- [18] S.-W. Wang, Y. Lan, and L.-H. Tang, J. Stat. Mech. Theor. Exp. **2015**, P07025 (2015).
- [19] C. Jia, Phys. Rev. E **93**, 052149 (2016).
- [20] I. Derényi and R. D. Astumian, Phys. Rev. E **59**, R6219 (1999).
- [21] T. Hondou and K. Sekimoto, Phys. Rev. E **62**, 6021 (2000).
- [22] M. Esposito, Phys. Rev. E **85**, 041125 (2012).
- [23] K. Kawaguchi and Y. Nakayama, Phys. Rev. E **88**, 022147 (2013).
- [24] S.-W. Wang, K. Kawaguchi, S.-i. Sasa, and L.-H. Tang, Phys. Rev. Lett. **117**, 070601 (2016).
- [25] R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).
- [26] P. Martin, A. Hudspeth, and F. Jülicher, Proc. Natl. Acad. Sci. U.S.A. **98**, 14380 (2001).
- [27] D. Mizuno, C. Tardin, C. Schmidt, and F. MacKintosh, Science **315**, 370 (2007).
- [28] S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, Phys. Rev. Lett. **104**, 198103 (2010).
- [29] L. F. Cugliandolo, J. Kurchan, and L. Peliti, Phys. Rev. E **55**, 3898 (1997).
- [30] E. Dieterich, J. Camunas-Soler, M. Ribezzi-Crivellari, U. Seifert, and F. Ritort, Nat. Phys. **11**, 971 (2015).
- [31] T. Harada and S.-i. Sasa, Phys. Rev. Lett. **95**, 130602 (2005).
- [32] T. Harada and S.-i. Sasa, Phys. Rev. E **73**, 026131 (2006).
- [33] S. Toyabe, H.-R. Jiang, T. Nakamura, Y. Murayama, and M. Sano, Phys. Rev. E **75**, 011122 (2007).
- [34] E. Lippiello, M. Baiesi, and A. Sarracino, Phys. Rev. Lett. **112**, 140602 (2014).
- [35] K. Sekimoto, Progr. Theor. Exp. Phys. **130**, 17 (1998).
- [36] G. Diezemann, Phys. Rev. E **72**, 011104 (2005).
- [37] C. Maes and B. Wynants, Markov Processes Relat. Fields, **45** (2010).
- [38] M. Baiesi, C. Maes, and B. Wynants, Phys. Rev. Lett. **103**, 010602 (2009).
- [39] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, Vol. 1 (Elsevier, 1992).
- [40] J. Sakurai and S. Tuan, *Modern Quantum Mechanics*, Advanced book program (Benjamin/Cummings Pub., 1985).
- [41] K. Kawaguchi, S.-i. Sasa, and T. Sagawa, Biophys. J. **106**, 2450 (2014).
- [42] S. Fielding and P. Sollich, Phys. Rev. Lett. **88**, 050603 (2002).
- [43] P. Sartori and Y. Tu, Phys. Rev. Lett. **115**, 118102 (2015).
- [44] J. Kurchan, Nature **433**, 222 (2005).
- [45] S. Toyabe and E. Muneyuki, New Journal of Physics **17**, 015008 (2015).
- [46] J. Tailleur and M. E. Cates, Phys. Rev. Lett. **100**, 218103 (2008).

## Appendix A: Real part of the response spectrum

Now, we prove that the following summation

$$\sum_{j=2}^N \alpha_j \phi_j \left[ 1 - \frac{1}{1 + (\omega/\lambda_j)^2} \right]$$

gives a real function over frequency domain, although each component can be a complex function.

For the first quantity, inserting the definition of the projection coefficients  $\alpha_j$  and  $\phi_j$ , we obtain

$$\begin{aligned} \sum_{j=2}^N \alpha_j \phi_j \left[ 1 - \frac{1}{1 + (\omega/\lambda_j)^2} \right] &= \sum_{j=1}^N \alpha_j \phi_j \left[ 1 - \frac{1}{1 + (\omega/\lambda_j)^2} \right] \\ &= \sum_{n,m} \mathcal{Q}_n \left( \sum_{j=1}^N x_j(n) \left[ 1 - \frac{\lambda_j^2}{\lambda_j^2 + \omega^2} \right] y_j(m) \right) B_m. \\ &= \sum_{n,m} \mathcal{Q}_n \left( 1 - \frac{M^2}{M^2 + \omega^2} \right)_{nm} B_m, \end{aligned} \tag{A1}$$

which indeed is a real function ( $M$  is a real matrix). Here,  $(\cdot)_{nm}$  takes the entry of the matrix at  $n$ -th row and  $m$ -th column. The above calculation has used a critical relation

$$\sum_{j=1}^N x_j(n) f(\lambda_j) y_j(m) = \left( f(M) \right)_{nm}, \tag{A2}$$

$$\sum_{j=1}^N x_j(n) f(\lambda_j) y_j(m) = \left( f(M) \right)_{nm}, \tag{A3}$$

where  $f(\cdot)$  is an analytical function. It can be proved as follows

$$\begin{aligned} \sum_{j=1}^N x_j(n) f(\lambda_j) y_j(m) &= \sum_{j=1}^N \sum_k \left( f(M) \right)_{nk} x_j(k) y_j(m) \\ &= \sum_k \left( f(M) \right)_{nk} \left[ \sum_{j=1}^N x_j(k) y_j(m) \right] \\ &= \sum_k \left( f(M) \right)_{nk} \delta_{km} \end{aligned}$$

Then summation over  $k$  gives the desired relation in Eq. (A3).

Similarly, we can show that

$$\sum_{j=2}^N \alpha_j \phi_j \left[ \frac{\omega/\lambda_j}{1 + (\omega/\lambda_j)^2} \right]$$

is a real function, and thus the real part of the response spectrum is given by

$$\tilde{R}'_{\dot{Q}}(\omega) = \sum_{j=2}^N \alpha_j \phi_j \left[ 1 - \frac{1}{1 + (\omega/\lambda_j)^2} \right]. \quad (\text{A4})$$

This justifies the violation spectrum in Eq. (28).

## Appendix B: The sum rule

Here, we prove the sum rule in Eq. (26). Using the definition  $\beta_j \equiv \sum_n y_j(n) P_n^{ss} \mathcal{Q}_n$  and the characteristic equation  $\lambda_j y_j(m) = -\sum_n y_j(n) M_{nm}$ , we find that

$$\lambda_j \beta_j = \sum_{n,m} y_j(n) \left[ w_n^m P_n^{ss} \mathcal{Q}_n - w_m^n P_m^{ss} \mathcal{Q}_m \right]. \quad (\text{B1})$$

Combined with  $\phi_j \equiv \sum_n B_n y_j(n)$  and Eq. (25), we obtain

$$\lambda_j \beta_j - T \phi_j = \sum_{n,m} y_j(n) (w_n^m P_n^{ss} - w_m^n P_m^{ss}) \frac{\mathcal{Q}_n + \mathcal{Q}_m}{2}. \quad (\text{B2})$$

Now, we add another component  $\alpha_j = \sum_{n'} x_j(n') \mathcal{Q}_{n'}$ , and check the summation over all eigenmodes:  $\sum_j \alpha_j (\lambda_j \beta_j - T \phi_j)$ . Noting the completeness relations  $\sum_j x_j(n') y_j(n) = \delta_{n',n}$ , we have

$$\sum_{j=1}^N \alpha_j (\lambda_j \beta_j - T \phi_j) = \frac{1}{2} \sum_{n,m} \mathcal{Q}_n (w_n^m P_n^{ss} - w_m^n P_m^{ss}) \mathcal{Q}_n + \frac{1}{2} \sum_{n,m} \mathcal{Q}_n (w_n^m P_n^{ss} - w_m^n P_m^{ss}) \mathcal{Q}_m.$$

On the right hand side (rhs), the stationary condition  $\sum_m [w_n^m P_n^{ss} - w_m^n P_m^{ss}] = 0$  demands that the first term vanishes, and the second term cancels the third term due to their equivalence under exchange of  $n$  and  $m$ . Then, we obtain

$$\sum_{j=1}^N \alpha_j (\lambda_j \beta_j - T \phi_j) = 0.$$

Noting  $\lambda_1 = 0$  and  $\phi_1 = 0$ , we have derived the sum rule in the main text.

### Appendix C: Detailed balance in each eigenmode for equilibrium system

We have considered a symmetric form of perturbation [Eq.( 5)] in the main text. Combining detailed balance condition

$$w_n^m P_n^{eq} = w_m^n P_m^{eq} \quad (C1)$$

and Eq. (B2) (note that  $ss$  should be replaced by  $eq$  since we assume that steady state is an equilibrium state), we can immediate obtain Eq. (30), the detailed balance in each eigenmode. However, Eq.( 5) is only a special case of the general form of perturbation:

$$\frac{\tilde{w}_m^n}{\tilde{w}_n^m} = \frac{w_m^n}{w_n^m} \exp\left(\frac{1}{T}[\mathcal{Q}_n - \mathcal{Q}_m]h\right). \quad (C2)$$

Here, we prove Eq. (30) holds under this general assumption.

According to Eq. (22), the flux fluctuation vector satisfies

$$\begin{aligned} B_n^{eq} &\equiv \lim_{h \rightarrow 0} \partial_h \sum_m \tilde{M}_{nm} P_m^{eq} \\ &= \lim_{h \rightarrow 0} \frac{\partial}{\partial h} \sum_m (\tilde{w}_m^n P_m^{eq} - \tilde{w}_n^m P_n^{eq}) \\ &= \lim_{h \rightarrow 0} \frac{\partial}{\partial h} \left( \sum_m \left[ \frac{\tilde{w}_m^n P_m^{eq}}{\tilde{w}_n^m P_n^{eq}} - 1 \right] \tilde{w}_n^m P_n^{eq} \right), \end{aligned} \quad (C3)$$

which holds for any form of perturbation. Using Eq. (C2) and  $P_n^{ss} = P_n^{eq} + O(h)$ , we have

$$\begin{aligned} \frac{\tilde{w}_m^n P_m^{eq}}{\tilde{w}_n^m P_n^{eq}} &= \frac{w_m^n P_m^{eq}}{w_n^m P_n^{eq}} \exp\left(\frac{1}{T}[\mathcal{Q}_n - \mathcal{Q}_m]h\right) \\ &= \frac{w_m^n P_m^{eq}}{w_n^m P_n^{eq}} \left(1 + \frac{1}{T}[\mathcal{Q}_n - \mathcal{Q}_m]h\right) + O(h^2) \\ &= 1 + \frac{1}{T}[\mathcal{Q}_n - \mathcal{Q}_m]h + O(h^2). \end{aligned}$$

To arrive at the last line, we have used the detailed balance condition Eq. (C1). Inserting this result into Eq. (C3) and taking the limit  $h \rightarrow 0$ , we obtain

$$B_n^{eq} = \frac{1}{T} \sum_m w_m^n P_m^{eq} (\mathcal{Q}_n - \mathcal{Q}_m), \quad (C4)$$

which is exactly the same as the equilibrium form of Eq. (25). Therefore,  $B_n^{eq}$  is independent of the specific form of perturbation and consequently Eq. (30) holds for each eigenmode.

### Appendix D: First order correction

Now, we consider the first order correction. For convenience, we introduce a compact form  $(g, f)$  as inner product between state function  $g(n)$  and  $f(n)$ , i.e.,  $(g, f) \equiv \sum_n g(n)f(n)$ . It is convenient to decompose  $x_j^{(1)}$  and  $y_j^{(1)}$  to the eigenspace of  $M^{(1)}$ , i.e.,

$$x_j^{(1)} = \sum_{j'} x_{j'}^{(0)} \left( y_{j'}^{(0)}, x_j^{(1)} \right), \quad (D1a)$$

$$y_j^{(1)} = \sum_{j'} y_{j'}^{(0)} \left( y_j^{(1)}, x_{j'}^{(0)} \right), \quad (D1b)$$

$$(D1c)$$

and then solve each component accordingly. By substituting into the  $O(1)$  order equation Eq. (??), and then projecting the resulting vector equations onto  $j'$ -th mode again, we easily obtain the first order correction for fast modes ( $j > K$ ), i.e.,

$$x_j^{(1)} = \sum_{j' \neq j} x_{j'}^{(0)} \frac{A_{j',j}}{\lambda_{j'}^{(-1)} - \lambda_j^{(-1)}}, \quad (\text{D2a})$$

$$y_j^{(1)} = \sum_{j' \neq j} y_{j'}^{(0)} \frac{A_{j,j'}}{\lambda_{j'}^{(-1)} - \lambda_j^{(-1)}}, \quad (\text{D2b})$$

where  $A_{j,j'} \equiv (y_j^{(0)}, M^{(0)} x_{j'}^{(0)})$ . The correction of eigenvalue  $\lambda^{(0)}$  for  $j$ -th fast mode is obtained by projecting Eq. (??) onto  $j$ -th fast mode, which gives ( $j > K$ )

$$\lambda_j^{(0)} = -A_{j,j}. \quad (\text{D3})$$

The treatment for the slow modes ( $j \leq K$ ) requires more care. Substituting Eq. (D1) into Eq. (??), we find that only the components along the fast modes matters, and we obtain the projection coefficient along the fast modes ( $j \leq K, j' > K$ ):

$$(y_{j'}^{(0)}, x_j^{(1)}) = \frac{A_{j',j}}{\lambda_{j'}^{(-1)}}.$$

$$(y_j^{(1)}, x_{j'}^{(0)}) = \frac{A_{j,j'}}{\lambda_{j'}^{(-1)}}.$$

The components along the slow modes can be obtained by considering equation of order  $\epsilon$

$$-\lambda_j^{(-1)} x^{(2)}(p_k) - \lambda_j^{(0)} x_j^{(1)}(p_k) - \lambda_j^{(1)} x_j^{(0)}(p_k) = \sum_l M_{kl}^p x_j^{(2)}(p_l) + \sum_{q_l} M_{p_k q_l}^{(0)} x^{(1)}(q_l) \quad (\text{D4a})$$

$$-\lambda_j^{(-1)} y^{(2)}(p_k) - \lambda_j^{(0)} y_j^{(1)}(p_k) - \lambda_j^{(1)} y_j^{(0)}(p_k) = \sum_l y_j^{(2)}(p_l) M_{lk}^p + \sum_{q_l} y^{(1)}(q_l) M_{q_l p_k}^{(0)}. \quad (\text{D4b})$$

Again, by substitution and projection, we obtain for  $j, j' \leq K$

$$(y_{j'}^{(0)}, x_j^{(1)}) = \frac{1}{\lambda_{j'}^{(0)} - \lambda_j^{(0)}} \sum_{j'' > K} A_{j',j''} (y_{j''}^{(0)}, x_j^{(1)}),$$

$$(y_j^{(1)}, x_{j'}^{(0)}) = \frac{1}{\lambda_{j'}^{(0)} - \lambda_j^{(0)}} \sum_{j'' > K} (y_j^{(1)}, x_{j''}^{(0)}) A_{j'',j'},$$

which turns out to be generated by first order correction in the fast modes. The correction of eigenvalue  $\lambda_j^{(1)}$  is given by projecting Eq. (D4) onto  $j$ -th slow mode, i.e., ( $j \leq K$ )

$$\lambda_j^{(1)} = -(y_j^{(0)}, M^{(0)} x_j^{(1)}) = - \sum_{j' > K} \frac{A_{j,j'} A_{j',j}}{\lambda_{j'}^{(-1)}}, \quad (\text{D5})$$

where only the fast modes matter. Inserting these coefficients back in Eq. (D1), we obtain the first order correction for the slow modes.